Dissipation and control in microscopic nonequilibrium systems

by

Steven J. Large

B.Sc., University of Guelph, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Physics
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2020

Declaration of Committee

Name:	**Steven J. Large**
Degree:	**Doctor of Philosophy**
Thesis title:	**Dissipation and control in microscopic nonequilibrium systems**
Committee:	**Chair:** Malcolm Kennett Associate Professor, Physics

David Sivak
Supervisor
Associate Professor, Physics

John Bechhoefer
Committee Member
Professor, Physics

Eldon Emberly
Committee Member
Professor, Physics

Nancy Forde
Examiner
Professor, Physics

Michael R. DeWeese
External Examiner
Associate Professor, Physics
University of California, Berkeley

Abstract

Quantifying the flow of energy, entropy, and information within and through nonequilibrium systems remains a central challenge in understanding the microscopic physics of biological systems. Over the past two and a half decades, parallel developments in the fields of theoretical stochastic thermodynamics and single-molecule experiments have made tremendous steps towards this end, advancing our understanding of the fundamental physical limitations and constraints faced by biological systems *in vivo*. Central in this focus are molecular machines: nanoscale protein complexes which interconvert between different forms of energy to perform useful functions to the cell. While single-molecule experiments on molecular machines have predicted impressively high efficiencies, much is still unknown about their performance *in vivo*. In this thesis we build upon these primitives, largely by making use of near-equilibrium phenomenological models to simplify and make tractable the problem of quantifying dissipation in molecular machines and predicting the operational modes which are imperative to minimizing their dissipation. By exploring the relevance of near-equilibrium models in the experimental investigation of a DNA hairpin, we find that such an approach can provide utility in understanding the strategies to reduce dissipation in nonequilibrium processes. However, single-molecule manipulations are significantly separated from the *in vivo* dynamics of molecular machines, and thus for the remainder of the thesis we expand upon this approach in various ways, generalizing the existing theoretical framework to more closely parallel the dynamics of molecular machines. By incorporating the inter-system feedback present in molecular machines, we find that familiar intuitions about how excess work and entropy production are related break down. Finally, we derive a phenomenological expression for the energy flows communicated within the components of a mechanochemical molecular machine. Ultimately, our analysis shows that intersystem feedback can lead to nonvanishing energy flows which are the manifestation of a Maxwell demon in the molecular machine itself.

Keywords: Nonequilibrium statistical physics; Biophysics; Molecular machines; Irreversible thermodynamics

– To the memory of *Thomas George Smith*

Acknowledgements

First and foremost, I would like to thank my supervisor David Sivak for his support and guidance over the past 5 years. You granted me independence to follow my interests and immerse myself in the truest essence of academic research. The experience I have gained by working with you has been essential in strengthening my capabilities, and has been vital to my continued development as a scientist.

Throughout my graduate education I have had the great opportunity of working closely with a number of collaborators, from whom I have benefited greatly. I would like to thank Carlos Bustamante, Shixin Liu, and particularly Sara Tafoya for their patience and perseverance working on the DNA hairpin project. I would also like to thank Raphaël Chetrite and Jannik Ehrich for their collaboration on various projects throughout the past few years, I greatly enjoyed working with you both.

I would also like to thank the guidance of my supervisory committee, John Bechhoefer and Eldon Emberly. Your unique approaches to scientific problems and continued feedback through the years has helped to broaden the scope of my thinking. The content of this thesis has been improved and clarified significantly from your encouragement and suggestions.

I greatly appreciate the financial support I have received throughout my graduate studies, including NSERC CGS-M and CGS-D3 fellowships, as well as the Billy Jones and Howard Malm graduate awards in physics, made possible through the continued generosity of SFU donors.

I would also like to thank my parents, Andrew and Sandra, for believing in me through the past 5 years and beyond. Ever since I left to start my undergraduate degree in 2011 at the University of Guelph, I could always depend on your encouragement and unwavering support. I also would like to thank my sister Jenny, her husband Ian, and son Nate, for always being there for me.

Through the ups and downs of graduate school, friends and colleagues make the worst of times better and the best of times great. I would like to acknowledge the entire cast of Sivak group members over the past five years, with a special mention of Aidan Brown and Alexandra Kasper, who helped immensely in the early years of grad school. I also benefited greatly from many colleagues and friends in the broader physics graduate student body, with a particular thanks owing to Chapin Korosec, Konstantin Lehmann, and Benny Jäger. I also want to thank the friends who have been along for the whole ride, Jordan Elvedahl

and Scott Matthews, such lasting friendships burn a mark on us all, your role in keeping me tethered to reality through the last decade cannot be understated.

Finally, I would like to thank my future wife Maegan Kelleway. Your support and encouragement throughout graduate school have been indispensable, ultimately, you—above all others—have helped me keep a clear head through this whole process. The past 5 years especially have brought many changes to our life, and this work wouldn't have been possible without you. This closes one chapter on our life and begins another, *one day* at a time.

The general struggle for existence is not a struggle for
raw materials—these for organisms are air, soil and water,
all abundantly available—nor for energy which exists
in plenty in the sun and any hot body in the form of heat,
but rather a struggle for entropy, which becomes available
through the transition of energy from the hot sun to the cold earth. [1]

– Ludwig Boltzmann

Table of Contents

List of Figures

Chapter 1

Introduction

The classical theory of thermodynamics deals with phenomenological relationships between flows of energy in macroscopic systems. While originally motivated by a need to understand and optimize the performance of steam engines in the 1800s [2], the scope of thermodynamics has expanded through the past two centuries, finding its place as a mathematical framework that pervades the branches of science and industry. At its heart, the theory of classical thermodynamics captures the mathematical relationships between physical quantities, such as heat and work, subject to a set of underlying assumptions. So long as these assumptions are valid, the classical theory of thermodynamics can be encompassed by four axiomatic rules known as the *laws of thermodynamics* [3].

Despite the widespread successes of classical thermodynamics, there are a number of shortcomings that limit its utility. Central to the focus of this thesis is the assumption of systems being in a state of equilibrium. The restriction of thermodynamic analysis to equilibrium precludes the treatment of many interesting situations, such as biological systems, which are manifestly out of equilibrium. Furthermore, through the subsequent development of a statistical theory of thermodynamics, known as statistical mechanics, it became clear that the laws of classical thermodynamics pertain to ensemble-average quantities.

For macroscopic systems, an appeal to the central limit theorem makes clear that for $N \sim 10^{23}$ interacting molecules, the fluctuations in energy and energy flows can be assumed to be vanishingly small (so long as the system is not sufficiently close to a critical point). However, for small systems—often on the scale of micro- or nano-meters—thermal fluctuations can no longer be ignored, and have significant implications for the overall physical description of the system of interest. Furthermore, ever since Clausius, in 1865, first articulated the law of increase in entropy [4], what has come to be known as the second law has posed a deep, unresolved question: how can microscopic equations of motion, that are symmetric under time reversal, give rise to macroscopic behavior that is not?

Over the past 25 years, the development of stochastic thermodynamics has ushered in a new era in thermodynamics, largely overcoming these limitations, and deepening our understanding of the second law as a statistical relation (Sec. 2.3). At its heart, stochastic

thermodynamics seeks to understand how the laws of thermodynamics manifest in microscopic strongly fluctuating systems that are potentially far from equilibrium. The theoretical framework of stochastic thermodynamics provides a consistent method of assigning physical quantities–such as work, heat, and entropy–to fluctuating systems in contact with thermodynamic reservoirs, even when those systems are far from equilibrium. These physical quantities can be identified along a single trajectory or at the level of probability distributions, thus permitting a diverse set of methods to understand the physics of thermodynamic systems across all scales.

Given a set of state energies, the equilibrium ensemble of a system can be determined without any knowledge of its dynamics. Thus, the entire theory of equilibrium statistical mechanics can be built without the need to model the microscopic dynamics. In nonequilibrium systems, however, the same is not true. In general, the state of a particular nonequilibrium system depends not only on its present conditions, but also on its history [5].

In addition to deriving a set of consistent laws of stochastic thermodynamics, the study of fluctuating systems led to the development of an entirely new class of results, known collectively as *fluctuation theorems* [6, 7, 8, 9, 10, 11, 12] (Sec. 2.4). These mathematical identities place stringent constraints on the fluctuations in stochastic systems, even far from thermodynamic equilibrium, and can be viewed as generalizations of the second law of thermodynamics. In particular, the entropy production fluctuation theorem can be used to derive the second law, and provides a much deeper understanding of its physical origins. In fact, all fluctuation theorems–even those not relating to entropy–give rise to second-law-like inequalities [13, 14]. This realization has led to a significant improvement in our understanding of the physical origins of irreversibility in nonequilibrium systems. For instance, the generalized Jarzynski equality [15] has shown how incorporating information—in the form of feedback control—into the thermodynamics of fluctuating systems can lead to subzero bounds on entropy production, a result which has been experimentally verified [16]. The incorporation of information into the theory of thermodynamics has effectively 'exorcised' the long-standing thought experiment known as *Maxwell's demon* [17]: by treating information as a physical quantity—which was suggested by Landauer in 1961 [18]—the familiar form of the second law is restored.

In tandem with the theoretical developments of stochastic thermodynamics, a new set of experimental techniques were developed to directly probe and perturb microscopic systems [19]. Early experiments were concerned with measuring the fluctuations in microscopic systems, and primarily aimed at verifying the fluctuation theorems. Examples include measurements of heat and work fluctuations in the tip of an AFM [20], the rotational angle of a torsion pendulum [21], and in an electrical resistor [22]. Furthermore, the development—and subsequent refinement—of optical tweezers for single-molecule force spectroscopy allowed the direct verification of theoretical predictions in biological systems, such as a DNA or RNA hairpin [23, 24, 25].

Biological systems have served as a central focus for applying stochastic thermodynamics. Ultimately, this comes as a result of the nonequilibrium nature of microscopic biological systems. The essential importance of nonequilibrium physics in our understanding of biology is summarized effectively by Erwin Schrödinger in *What is Life?* where he famously equated the 'decay into thermodynamic equilibrium' with death [26]. More recently, a great deal of effort has focused on our understanding of molecular machines, which are a class of nanoscale objects that operate out of equilibrium to perform useful tasks within biological cells. These molecular machines manage to remain out of thermodynamic equilibrium by siphoning free energy off of nonequilibrium environmental conditions, such as out-of-equilibrium concentrations of chemical reactants and products. In fact, the understanding of molecular machines has served as a central motivation for the continued development of stochastic thermodynamics as a whole [27, 28].

Stochastic thermodynamics can help in furthering our understanding of the physics of molecular machines, in particular, the fundamental physical constraints such nonequilibrium systems face *in vivo*, and how these constraints affect the limits of performance in molecular machines. Thus, there is a need to define a performance metric in such systems that quantifies the loss of capacity in the system. The entropy production, or dissipation, can serve such a role: entropy production represents a fundamental loss of system capacity to perform useful work. Low-entropy states have a higher capacity to perform work than high-entropy states; for instance a fully extended polymer chain has the capacity to exert forces during its compaction. Many recent significant advances, such as the thermodynamic uncertainty relation [29, 30] and its generalizations [31, 32, 33, 34], have been motivated as providing such fundamental functional limitations on the performance of molecular machines out of equilibrium.

1.1 Molecular machines

At the sub-cellular scale, biological systems exhibit a strikingly high degree of organization which is inconsistent with an equilibrium state [27]. This organization in maintained, in large part, through the concerted effort of a host of molecular machines [35]. These nanoscale machines consume energy, typically in the form of high-energy chemical bonds, to perform useful functions in the cell [28].

Physically, molecular machines are made up of several interacting soft-matter components, which experience large thermal fluctuations. Over the past several decades, molecular machines have been a subject of intense focus within the biophysics community, and have been studied using a host of modern experimental techniques, such as optical tweezers [36, 37, 38, 39], or electrorotation [16, 40, 41, 42], to probe and perturb individual molecular machines. Efforts to better understand the physics of molecular machines promise to deepen our understanding of the fundamental operational constraints facing evolved molec-

ular machines, but also promise a range of practical benefits, such as the design and synthesis of *de novo* molecular machines, perhaps accelerating their use in next-generation nanomedicine [43, 44].

1.1.1 Kinesin

A canonical example of a molecular transport motor is kinesin [45]. The first of many variants, kinesin-1, was discovered in 1985 in observations of the mobility of organelles in the extruded cytoplasm from the giant axon of a squid [46, 47, 48]. Soon after, this 'axoplasmic motility' was understood to be due to the effect of 'a novel force-generating molecule' which has since come to be known as kinesin-1 [49]. Many distinct variations of the kinesin motor were found soon after by identifying a common genetic motif in the genome of *Drosophila melanogaster* [50], and subsequently found to occur, in one form or another, in virtually all forms of eukaryotic cells [51].

Kinesin is a molecular motor that is responsible for the transport of cellular cargoes throughout individual cells. Morphologically, kinesin consists of two motor head domains, which are connected through a long stalk to two cargo binding domains. By making use of the high-energy phosphate bond in the cellular energy currency ATP, kinesin moves by its head domains processively walking along a microtubule, typically in a '+'-end directed fashion, although certain kinesin motors can, in fact, travel in the opposite direction [52]. *In vivo*, the cell ensures directionality by maintaining an out-of-equilibrium concentration of ATP and ADP molecules [45, 51].

Through the use of modern experimental techniques, kinesin has been studied in great detail, informing much of our modern understanding of the mechanics of nanoscale machines. For instance, early studies with optical trapping allowed direct measurement of the stall force—the maximum opposing force under which the kinesin can still move in the forward direction—of $\sim$ 6–8 pN [36, 37, 38, 39]. When this stall force is multiplied by the kinesin step size of $\sim$ 8nm, the resulting work done by the kinesin motor is $\sim$ 48–64 pN$\cdot$ nm, which, when compared with the free energy $\sim$ 80 pN$\cdot$nm liberated by hydrolysis of a single ATP, indicates an efficiency of up to $\sim$ 80% [53]. Additionally, single-particle fluorescence tracking has measured the *in vivo* speeds of kinesin motors directly, clocking in at a maximum of $\sim$ 1 µm/sec, or 125 (8-nm) steps/sec [54].

In addition to providing an interesting biophysical model for understanding the physics of molecular machines, kinesins are suspected to play a central role in pathophysiology of many diseases. In particular, their role in transport implicates them in many neuronal disorders, such as amyotrophic lateral sclerosis (ALS) and Alzheimer's, where neurons' extremely polarized cell morphologies make processive transport motors—such as kinesin—vital to their healthy function [55, 56].

1.1.2 ATP synthase

With ATP serving as the primary chemical fuel source for many molecular machines (such as kinesin), the production of ATP is essential to the operation of biological cells. The molecular machine ATP synthase is responsible, in large part, for the production and maintenance of cellular ATP stores. As with kinesin, several variations of ATP synthase exist in different organisms, but its presence is ubiquitous amongst both eukaryotic and prokaryotic cells [57, 58]. In eukaryotic cells, ATP synthase motors are often located in the membranes of organelles, whereas in prokaryotic cells they are found in the cell membrane. In fact, the internalization of ATP synthase—and corresponding increase in cellular energy capacity—during the emergence of eukaryotic cells in the evolutionary past is thought to have been a central factor in the subsequent formation of higher-order structures, such as complex multicellular organisms [59].

While there are several different forms of ATP synthase, the most well-studied is F_oF_1 ATP synthase, which is found in eukaryotic mitochondria and chloroplasts as well as prokaryotic cell membranes. In contrast, variants such as the V_oV_1 ATPase inhabit organelle membranes other than mitochondria, such as endosomes and lysosomes [60, 61], while A_oA_1 ATP synthase appears in the membrane of extremophilic archaea. V_oV_1 ATPase operates in the opposite direction of F_oF_1, using ATP as an energy source to acidify the interior of organelles, and contains a biochemical control mechanism, whereby the depletion of cellular glucose causes dissociation of the V_o and V_1 components, inactivating the complex as an ATPase [61, 62]. Even within the class of F_oF_1 ATP synthase motors, however, there is significant diversity across organisms. For instance, the number of protein subunits in the F_o component can vary between organisms, and can have a significant impact on its overall energetics [63].

The remarkable efficiencies achieved by ATP synthase make it a fascinating example of energy transduction in out-of-equilibrium biological systems. The ATP synthase motor provides a straightforward example of a molecular machine that converts energy stored in out-of-equilibrium chemical concentrations ($[H^+]$ difference across the membrane) into an essential cellular resource (ATP). This has made ATP synthase a canonical system in which to model and understand the fundamental physics of molecular machines and the inherent trade-offs in the functional capabilities of biomolecular machines [64, 65, 66, 67].

Throughout this thesis, the simple models we use to test theoretical models are motivated by the physics of molecular machines, such as kinesin and ATP synthase. These represent a minute fraction of the diversity of molecular machines that exist within biological cells, performing a wide array of functions, and largely responsible for maintaining the nonequilibrium conditions necessary for life. Fundamentally, these nanoscale engines operate in the presence of strong thermal fluctuations and out of thermodynamic equilibrium,

and thus to understand their physics we need to make use of nonequilibrium theories of statistical physics and stochastic processes.

1.2 Nonequilibrium statistical physics

For a physical system in contact with a thermal reservoir, at equilibrium the microstates x are distributed in accordance with the Boltzmann distribution

$$\pi(x|\boldsymbol{\lambda}) = e^{-\beta E(x|\boldsymbol{\lambda})+\beta F(\boldsymbol{\lambda})}, \tag{1.1}$$

where $\beta \equiv (k_{\mathrm{B}}T)^{-1}$ is the inverse temperature of the reservoir, $E(x|\boldsymbol{\lambda})$ is the energy of microstate x given a set of external parameters $\boldsymbol{\lambda}$—such as pressure or volume—and $F(\boldsymbol{\lambda})$ is the equilibrium free energy of the system. The equilibrium free energy is related to the partition function by $F(\boldsymbol{\lambda}) \equiv -k_{\mathrm{B}}T \ln \mathcal{Z}$, where $\mathcal{Z} \equiv \int \exp\left[-\beta E(x|\boldsymbol{\lambda})\right] \mathrm{d}x$ normalizes the equilibrium distribution. Thus, to determine the equilibrium distribution for any system of interest, we only need to know the energies of each of the microstates.

In contrast, determining the nonequilibrium distribution of a system requires knowledge of its dynamics. Furthermore, aside from special cases, like a nonequilibrium steady state, determining the instantaneous nonequilibrium distribution over microstates also requires knowledge of the previous history of the system. For systems that we are generally interested in, such as a polymer in solution, there are a very large number of individual molecules, $N \approx 10^{23}$ (counting all of the waters) that can have non-negligible interactions with the system. The explicit modeling of all degrees of freedom in such a system is analytically intractable, and even outside of the realm of numerical calculations.

Fortunately, it is often the case that a relatively small number of these molecules represent our system of interest. Thus, we seek a method of obtaining equations of motion for the system of interest that do not require explicitly accounting for all the molecules that make up the environment. Mathematically, such reduced descriptions of the system dynamics can be justified by an assumption of weak interactions between the system and its surroundings, or a separation of timescales between the system and environment. Weak interactions between the system and its environment are a primary assumption behind, for instance, Zwanzig's projection operator method [5], where the Liouville equation for the entire system is projected onto a low-dimensional space that captures the dynamics of the system of interest, and the effective interactions between the system and environment are captured through an additional fluctuating force. Conversely, the assumption of a separation of time scales enforces a typical axiom of irreversible thermodynamics: thermodynamic reservoirs always remain at equilibrium [68]. In either case, the same equations of motion are the result, which transform the typical differential equations of classical mechanics to stochastic differential equations, with fluctuating components that capture the effective interactions between a system and its surroundings.

A canonical example for the equation of motion of a stochastic system is the diffusive motion of a small object, such as a pollen grain, in water. This phenomenon was first documented by the botanist Robert Brown in 1827 while observing the motion of pollen grains in water [69], however it wasn't until the early 20th century when Einstein [70]—and, independently, Smoluchowski [71]—provided a clear and consistent interpretation of the seemingly random dynamics. Einstein proposed that the random motion of the pollen grain was due to its repeated interaction with the water molecules surrounding it. Furthermore, because the motion of all of the individual water molecules was exceedingly complex, the motion of the pollen grain could only be understood statistically. As such, Einstein's treatment of what is now called Brownian motion is often seen as the first treatment of a physical phenomenon using stochastic modeling [72].

These two main points led Einstein to the famous *diffusion equation*, quantifying the temporal evolution of the probability of observing the position of the pollen grain at a particular point in space

$$\frac{\partial p(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)}{\partial t} = D\nabla^2 p(\boldsymbol{x}, t|\boldsymbol{x_0}, t_0) \ , \tag{1.2}$$

where $p(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)$ is the probability density of observing the pollen grain at position $\boldsymbol{x}$ at time t given it was at $\boldsymbol{x}_0$ at time t_0, ∇^2 is the Laplacian operator, and D is the diffusion coefficient. The diffusion coefficient D is related to the rate at which the probability distribution spreads out in time. Specifically, the mean-squared displacement (MSD) of a diffusive particle obeys the relationship

$$\langle \delta x^2(t) \rangle = 2N_{\text{dim}} Dt \ , \tag{1.3}$$

where the angle brackets $\langle \cdots \rangle$ indicate an average over an ensemble of pollen grains, all initialized at the same position at time $t = 0$, $\delta x \equiv x - \langle x \rangle$ is the deviation of x from its average position, and N_{dim} is the number of dimensions in which the particle is diffusing. For a pollen grain diffusing on the surface of water, $N_{\text{dim}} = 2$.[1]

Not long after Einstein's discovery of the diffusion equation, Paul Langevin, a French physicist, presented an alternative treatment of the process of diffusion using stochastic processes. Specifically, he claimed that the equation of motion of an individual pollen grain, as per Newtonian mechanics, is given by the differential equation

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\gamma \frac{\mathrm{d}x}{\mathrm{d}t} + F(t) \tag{1.4}$$

[1]There are, however, some subtleties involved in diffusion of a particle at an air-water interface. Specifically, the diffusion coefficient predicted by the Einstein relation (1.7) may be modified due to the interaction of the particle with the interface. For instance, an increase in drag forces due to surface tension [73], the Marangoni effect and capillary forces [74], and even the elastic response of the interface to fluctuation-induced surface-wave formation [75] can have significant impact on the value of the diffusion coefficient.

where the first term on the RHS represents a frictional damping force $F_{\text{fric}}(t) = -\gamma v(t)$, with γ the viscous damping coefficient for the pollen grain in the fluid, and $F(t)$ is a random variable that represents a fluctuating force, arising from the frequent impacts of water molecules with the pollen grain. While it would take another 40 years to construct a mathematically rigorous derivation of the form of the fluctuating force, we will simply state two properties of the fluctuating force that turn out to be true:

$$\langle F(t) \rangle = 0 \tag{1.5a}$$

$$\langle \delta F(t) \delta F(t') \rangle = 2 k_{\text{B}} T \gamma \delta(t - t') \ . \tag{1.5b}$$

The first of these properties (1.5a) implies that the fluctuating force does not, on average, impart any net force on the pollen grain, while the second (1.5b) suggests that the random forces at different times are uncorrelated. The constant of proportionality in (1.5b) is fixed by the equipartition theorem [5].

Using (1.5a) and (1.5b), we can write the Langevin equation in its more familiar form

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t} = -\gamma \frac{\mathrm{d}x}{\mathrm{d}t} + \sqrt{2 k_{\text{B}} T \gamma}\, \xi(t) \ , \tag{1.6}$$

where now $\xi(t)$ is a zero-mean $\langle \xi(t) \rangle = 0$, delta-correlated *white noise process* with $\langle \xi(t) \xi(t') \rangle = \delta(t - t')$. Finally, we can connect the phenomenological constants γ from the Langevin equation (1.6) and D from the diffusion equation (1.2) through the *Einstein relation*

$$D = \frac{k_{\text{B}} T}{\gamma} \ . \tag{1.7}$$

Fundamentally, the diffusion equation (1.2) and Langevin equation (1.6) represent two different approaches to solving the same problem: what is the equation of motion for a fluctuating particle in contact with a thermal reservoir? In each case, the influence of the thermal reservoir comes into the mathematical description of the system through a transport coefficient (γ or D). The diffusion equation solves this problem by calculating the full probability distribution function $p(x, t)$ as the solution of the partial differential equation in (1.2), whereas the Langevin equation provides a description in terms of a stochastic differential equation, where the solution is a fluctuating trajectory consistent with the statistical properties of the dynamics.

Figure 1.1 shows a comparison between trajectory-level and probability distribution descriptions of a stochastic system. For $n = 1000$ Langevin trajectories (left), the empirical distribution over positions x at specific times match the solution to the corresponding (right) diffusion equation (1.2).

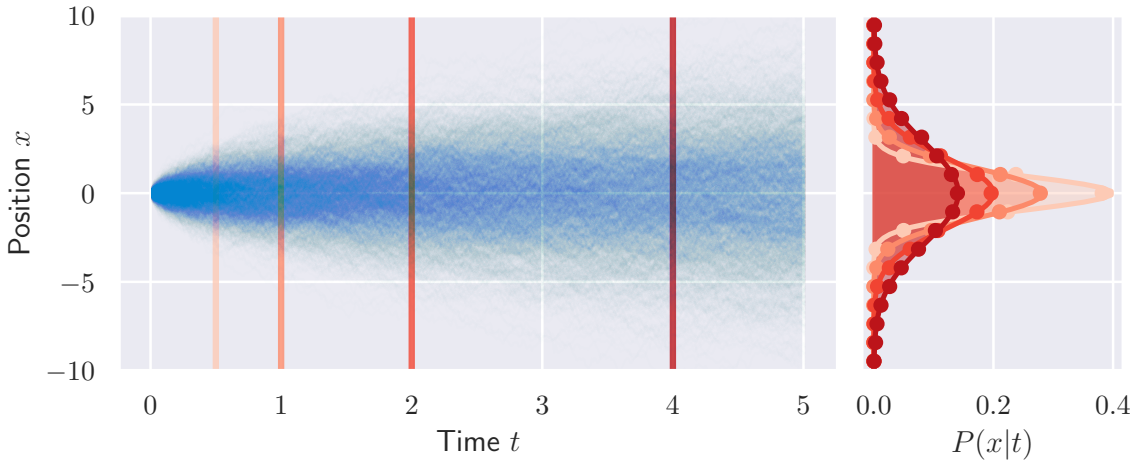

Figure 1.1: **Trajectory-level and distribution-level simulations of diffusive motion give alternative descriptions of the same physical process.** (left) A sample of $n = 1000$ individual trajectories of a Langevin simulation of Brownian motion compared to the (right) time-dependent probability distribution obtained by solving the diffusion equation (solid curves). Circles indicate the empirical distribution of positions obtained from the ensemble of Langevin trajectories at the same times (indicated by the red vertical bars on the left), showing agreement.

1.3 Overview of this thesis

The remainder of this thesis consists of three parts, discussing research that broadly falls into three categories: experimental tests of nonequilibrium theory (Part I), dissipation through the lens of control theory (Part II), and the nonequilibrium physics of autonomous machines (Part III). The content for Part I presents and elaborates on results published in [25], Part II discusses the theoretical work presented in [76, 77] and elaborates upon previous discussions in [76] on lower bounds on dissipation (Chapter 8). Finally, in Part III we discuss the implications of control-theoretical models of nonequilibrium physics in understanding the microscopic physics of molecular machines. Specifically, we investigate the relationship between excess work and entropy production in strongly coupled nonequilibrium system from [78], and derive a near-equilibrium phenomenological method of quantifying the energy flows between the components of such systems presented in [79].

1.4 Contributions to this thesis

Several chapters in this thesis contain results from collaborative efforts. The content in Part I (Ch. 3-5) is drawn from material published in [25] regarding experiments performed in collaboration with Sara Tafoya, Shixin Liu, and Carlos Bustamante at the University of California, Berkeley. In this project, Sara Tafoya performed the experiments and gathered

the raw data, and all parties in the collaboration participated in conceptualization of the project. I performed the data curation and formal analysis of the resulting data, designed experiments, and wrote the necessary software for the project, including both the data analysis and visualization. In summarizing the results and discussing how to present them, all parties in the collaboration took part.

Furthermore, Raphaël Chetrite helped with the formal analysis and mathematical details of our derivations in Chapter 6 (which are published in [76]), particularly for the exact solutions found in Sec. 6.4.1. Jannik Ehrich helped with conceptualization of the project discussed in Chapter 9 and published in [78], performed some of the formal analysis, and helped in writing up the results.

Chapter 2

Theoretical background

2.1 Mathematical preliminaries

We now develop the mathematical framework behind the stochastic equations of motion governing microscopic, fluctuating systems. In particular we will derive the three primary methods of describing stochastic dynamics: the master equation, the Langevin equation, and the Fokker-Planck equation. Each of these descriptions of stochastic dynamics is used throughout the remainder of the thesis. The master equation is used in Chapter 9, the Langevin equation is used in Chapters 6 and 10, and the Fokker-Planck equation is used in Chapter 7.

2.1.1 Random variables, probability distributions and characteristic functions

The fundamental mathematical underpinning of stochastic thermodynamics and nonequilibrium physics is probability theory. As a mathematical theory originally axiomatized by Andrey Kolmogorov in the 1930s [80], probability theory has become central to a number of fields of science and industry, from the fluctuating proteins within a biological cell to the statistics of returns in an investment portfolio [72, 81]. At its core, probability theory deals with the mathematics of uncertainty, quantifying the likelihoods of random events.

The outcome of a random event X, such as a coin toss, is referred to as a *random variable* because its outcome is uncertain. While we do not know what the exact outcome of the event will be, we know what the possible outcomes x may be, and can assign relative likelihoods to them. For instance, in tossing a fair coin, the possible outcomes are heads or tails, and both are equally likely. The function that quantifies the relative likelihoods for different outcomes of the random variable X is called the probability p_x, where here p_x is the likelihood that the outcome of the random event will be $X = x$. While a full treatise on the theory of probability is beyond the scope of this introduction[1], a central axiom of

[1]For the interested reader, however, [82] provides a succinct introduction to the mathematical subject.

probability theory is that the probability distribution is *normalized*,

$$\sum_x p_x = 1 \ , \tag{2.1}$$

where the summation is taken over all possible outcomes of the random event.

As defined in (2.1) the probability distribution is applicable to discrete random variables, and often referred to as the *probability mass function*. However, when the outcome of a random variable can be any value in a continuous domain, we instead use the *probability density function* $p(x)$, which satisfies the alternative normalization condition

$$\int_x p(x)\mathrm{d}x = 1 \ . \tag{2.2}$$

Intuitively, the need for a generalized definition of the probability in continuous spaces derives from the fact that there is no way of assigning the probability of a single outcome $X = x$ when x has a continuous domain, as each individual outcome has zero probability.[2] Instead, we must refer to the probability that the outcome x lies within a range of possible values $p(a < x < b) = \int_a^b p(x)\mathrm{d}x$ [81]. For the majority of this thesis, we will work with the probability density function, as most model systems we look at are continuous, but will distinguish between the two distributions with the notation used above: in probability mass functions the outcome appears as a subscript p_x, while for probability density functions the outcome appears as an argument $p(x)$.

While a mathematically complete description of a random variable requires specification of the probability distribution, we can describe particular properties of the random variable through its *moments*. The nth moment of the distribution is

$$\langle X^n \rangle = \int_x x^n p(x)\mathrm{d}x \ , \tag{2.3}$$

where here the angle brackets $\langle \cdots \rangle$ indicate an average over the distribution of x. Here, the distribution being averaged over is unambiguous, however in cases where it is not, the averaging distribution will be made explicitly clear from the notation, often with a subscript or superscript.

From (2.3), the zeroth moment ($n = 0$) of any distribution simply gives the normalization condition (2.2), and the first moment ($n = 1$) is the mean of the distribution, which is also often referred to as the *expectation value* of X. Higher-order moments encode more detailed information about the shape of the distribution, with each successively higher-order n providing information about events that are farther from the mean [72]. However, for $n > 1$ there are alternative, more convenient measures used to quantify the characteristics

[2] Or, in the language of probability theory, each individual event represents a set of zero measure.

of the distribution. For instance, the centered moments subtract the mean (first moment) to give moments relative to the mean value,

$$\langle \delta X^n \rangle = \int_x (x - \langle x \rangle)^n p(x) \mathrm{d}x \ , \tag{2.4}$$

where here, and in the remainder of the thesis, $\delta X \equiv X - \langle X \rangle$ indicates the difference between a random variable and its mean value. Trivially $\langle \delta X \rangle = 0$, and thus the first nonzero centered moment is for $n = 2$. $\langle \delta X^2 \rangle$ is the *variance*, which is often written as σ_x^2 as it is equal to the square of the standard deviation σ_x [83]. More convenient still are the *cumulants* κ_n of a distribution (see (2.7) below). The first two cumulants are simply the mean (κ_1) and variance (κ_2), and, while higher-order cumulants can be expressed as combinations of moments (2.3), no simple closed-form expression is known to do so [84].

An alternative route to understanding properties of a probability distribution comes through its Fourier transform, which is known as the *characteristic function* (or moment-generating function)

$$\phi(s) \equiv \langle e^{-isx} \rangle = \int_x e^{-isx} p(x) \mathrm{d}x \ . \tag{2.5}$$

The Fourier transform is used extensively in physics, particular when considering periodic behavior, such as waves [85]. In the context of probability theory, the characteristic function provides an often useful tool in solving problems, as well as a route to more simply calculate the moments (2.3) and cumulants of a distribution. In particular, the nth moment of a distribution can be calculated from the characteristic function as

$$\langle X^n \rangle = i^{-n} \frac{\partial^n \phi(s)}{\partial s^n} \bigg|_{s=0} \ . \tag{2.6}$$

Furthermore, the characteristic function can be used to define, mathematically, the cumulants through the *cumulant-generating function* $g(s) \equiv -\ln \phi(s)$. The cumulants κ^n can be determined from the cumulant-generating function, similar to (2.6), as

$$\kappa_n \equiv i^{-n} \frac{\partial^n g(s)}{\partial s^n} \bigg|_{s=0} \ . \tag{2.7}$$

The cumulant-generating function has become central in the field of large deviation theory, where a scaled cumulant-generating function is often used to determine the large deviation rate function [86]. However, the cumulants also have fundamental importance in standard treatments of probability theory. For instance, all cumulants above κ_2 (i.e., third order and higher) are zero for a Gaussian distribution, while no analogous truncations exist for the moments [84].

For situations described by multiple random variables, such as the outcome of several coin flips, the probability distribution of outcomes is multivariate. Here, there are several other important concepts to understand. In particular, for two random variables X and Y,

the probability of a particular outcome is now given by the *joint distribution* $p(x, y)$, which is the probability that both $X = x$ and $Y = y$. The joint distribution obeys the normalization condition $1 = \int_{x,y} p(x, y)\mathrm{d}x\mathrm{d}y$. The univariate distribution $p(x)$—also called the *marginal distribution* in this context—can be obtained by integrating—or marginalizing—over the y-coordinate as

$$p(x) = \int_y p(x, y)\mathrm{d}y \ . \tag{2.8}$$

Alternatively, given we know the outcome of Y, the probability of X conditioned on the outcome $Y = y$ is known as the *conditional distribution* $p(x|y)$.

We can relate all three distributions through the law of total probability as [83]

$$p(x, y) = p(x|y)p(y) = p(y|x)p(x) \ . \tag{2.9}$$

Finally, if the random variables X and Y are independent of one another, then the conditional and marginal distributions are equal ($p(x|y) = p(x)$ and $p(y|x) = p(y)$), and therefore the joint distribution can be expressed as a product of the marginal distributions

$$p(x, y) = p(x)p(y) \ . \tag{2.10}$$

2.2 Nonequilibrium dynamics

To describe the state of a fluctuating system, such as the position of a diffusing particle, we can gain tremendous insight by making use of the machinery of probability theory. In particular, the state of a system at a time t is itself a random variable $X(t)$ that depends on time. For instance, if we observe the state of the system at a series of times $t_0, t_1, \cdots, t_n$, so that increasing indices indicate observations that are later in time ($i > j$ implies that $t_i > t_j$), and each successive time is spaced with the increment $t_{i+1} - t_i = \Delta t$, then the system trajectory is a vector of random variables $X(t) = (X_n, X_{n-1}, \cdots, X_0) \equiv X_{t_{[0,n]}}$. Here, X_i is the state X at time t_i, and the subscript $t_{[0,n]}$ is shorthand notation for the sequence of times $t_0, t_1, \cdots, t_n$. This process is an example of a discrete-time stochastic process, because the random variable is observed at discrete time increments; however, this can be turned into a continuous-time process by taking the limit of $\Delta t \to 0$.

The probability of observing a particular trajectory of the system is given by the multivariate joint distribution $p(x_{[0,n]})$. Furthermore, by viewing the trajectory $x_{[0,n]}$ as a time series, the conditional distribution can be understood as a predictive tool, quantifying the probability distribution of future states of the system, given a previous history. For instance,

$$p(x_n|x_{[0,n-1]}) \equiv \frac{p(x_n, x_{[0,n-1]})}{p(x_{[0,n-1]})} \tag{2.11}$$

defines the probability of observing the system in state x_n given that the n previous observations were $x_{[0,n-1]}$.

An assumption that greatly simplifies the analysis of stochastic systems is the *Markov assumption*, which asserts that the state of the system at time t_n only depends on the most recent previous state, at t_{n-1}, and is thus conditionally independent of $x_{[0,n-2]}$. More concretely, the Markov assumption states that

$$p(x_n|x_{n-1}, x_{[0,n-2]}) = p(x_n|x_{n-1}) \ , \tag{2.12}$$

and systems which obey this assumption are said to be *Markovian*. The single-timestep conditional probabilities $p(x_{i+1}|x_i)$ are often called transition probabilities, as they represent the probability of observing the system in state x_{i+1} at time t_{i+1} given that it was in state x_i at time t_i. While there are many other common and useful assumptions that can be made on the interdependence of system states along a stochastic trajectory, such as the Martingale property that has become central to many formal treatments of stochastic processes and probability theory [87], the Markov assumption is, by and large, the most ubiquitous in the modern treatment of stochastic thermodynamics.

The primary utility of the Markov assumption is that it allows the decomposition of stochastic trajectories into transition probabilities. In particular, for a system which at time t_0 is distributed over x states as $p_{\text{init}}(x_0)$, the trajectory probability distribution over states after n time increments can be decomposed into a product of transition probabilities as

$$p(x_{[0,n]}) = p(x_n|x_{n-1})p(x_{n-1}|x_{n-2}) \cdots p(x_1|x_0)p_{\text{init}}(x_0) \ . \tag{2.13}$$

For a Markovian system, we can use (2.13) to write the trajectory probability $p(x_2, x_1, x_0) = p(x_2|x_1)p(x_1|x_0)p(x_0)$. By then marginalizing over x_1 (2.8) we obtain a tremendously useful identity in the study of stochastic processes called the *Chapman-Kolmogorov* equation:

$$p(x_2|x_0) = \int p(x_2|x_1)p(x_1|x_0)\mathrm{d}x_1 \tag{2.14}$$

for a continuous state space, or, for a discrete state space:

$$p(x_2|x_0) = \sum_{x_1} p(x_2|x_1)p(x_1|x_0) \ . \tag{2.15}$$

Under certain constraints on the limits of the transition probabilities as the time increment $\Delta t \to 0$, the Chapman-Kolmogorov equation (2.14) (or (2.15)) can be written as a

differential equation.[3] Specifically, for the transition rate $R(x|x',t) \equiv \lim_{\Delta t \to 0} \frac{p(x,t+\Delta t|x',t)}{\Delta t}$, one can show that the *differential Chapman-Kolmogorov equation* is [72]

$$\partial_t p(\boldsymbol{x},t|\boldsymbol{x}',t') = -\sum_i \partial_{x^i}\left[A_i(\boldsymbol{x},t)p(\boldsymbol{x},t|\boldsymbol{x}',t')\right] + \frac{1}{2}\sum_{i,j}\partial^2_{\boldsymbol{x}^i,\boldsymbol{x}^j}\left[B_{ij}(\boldsymbol{x},t)p(\boldsymbol{x},t|\boldsymbol{x}',t')\right] + \int \left[R(\boldsymbol{x}|\boldsymbol{x}',t)p(\boldsymbol{x}',t|\boldsymbol{x}',t') - R(\boldsymbol{x}|\boldsymbol{x}',t)p(\boldsymbol{x}',t|\boldsymbol{x}',t')\right] \mathrm{d}\boldsymbol{x} \ . \tag{2.16}$$

Here the state $\boldsymbol{x}$ is, in general, a vector, $\partial_t \equiv \partial/\partial t$ is the partial derivative with respect to time, and $\partial_{x^i} \equiv \partial/\partial x^i$ is the partial derivative with respect to the ith component of $\boldsymbol{x}$. This form indicates three different types of motion that can arise in stochastic dynamics. In particular, $A_i(\boldsymbol{x},t)$ represents a *drift* term, quantifying deterministic motion of the system, while $B_{ji}(\boldsymbol{x},t)$ is a positive-semidefinite diffusion tensor, quantifying the impact of diffusive motion. The final term on the RHS of (2.16) represents the impact of discrete jumps on the overall stochastic dynamics of the system. In various limits, the differential Chapman-Komogorov equation (2.16) reduces to common equations of motion for stochastic systems.

2.2.1 Master equation

For systems in which the drift and diffusion coefficients are zero ($A_i(\boldsymbol{x},t) = B_{ij}(\boldsymbol{x},t) = 0$, for all i,j), the resulting differential equation is the master equation, which describes the rate of change of the probability in a system where the source of probability changes are solely due to jumps. While one can use the master equation to describe the dynamics on continuous state spaces [72], for the purposes of this thesis, we will use the master equation when discussing discrete-state systems. The discrete master equation is often used to model the dynamics of chemical reaction networks [88], or as a simple representation of a more complex problem, for instance following coarse-graining [68].

Thus, the discrete master equation quantifies the rate of change of the probability distribution in a stochastic system due to jumps between discrete states. Mathematically, the discrete master equation is

$$\mathrm{d}_t p_i = \sum_j R_{ij} p_j \ , \tag{2.17}$$

where p_i is the probability of the system being in state i at time t, $\mathrm{d}_t \equiv \mathrm{d}/\mathrm{d}t$ is shorthand for the time derivative, and R_{ij} is the (i,j) element of the transition rate matrix, quantifying the rate of transitions from state $j \to i$ per unit time. R_{ij} is shorthand for $R(x_i|x_j)$ (here assumed to be time-independent), and the indices indicate that the state space is discrete. In order to conserve probability, the diagonal elements of the transition rate matrix are

[3] See Gardiner [72] for a more in-depth discussion of these constraints, but simply put, the constraints relate to the continuity of the underlying stochastic process in the continuous-time limit.

constrained by

$$R_{ii} = -\sum_{j \neq i} R_{ij} \ . \tag{2.18}$$

It can be shown, in several different ways, that in the long-time limit—and for time-independent transition rates—solutions to the master equation converge to a stationary distribution [84]. Furthermore, this distribution is unique if the transition rate matrix is irreducible, which means that each state in the system is accessible—via intermediate states—by any other state. Concretely, it is possible for a system initially in state x_i at $t = 0$ to be found in any other state of the system x_j at a later time. In what follows we will assume that there is a unique stationary solution, and will refer to it in general as the steady-state distribution. Once the dynamics have a thermodynamic interpretation, under certain circumstances the steady-state distribution is equivalent to the Boltzmann equilibrium distribution (1.1).

At steady state, the probability distribution is unchanging: $\mathrm{d}_t p_i = 0$. Therefore the stationary distribution p_i^{ss} solves the equation

$$\sum_j R_{ij} p_j^{\mathrm{ss}} = 0 \ . \tag{2.19}$$

Thus the stationary distribution is the (unique) right eigenvector of the transition rate matrix corresponding to the zero eigenvalue [84]. Thus for a time-independent rate matrix, one only needs to solve for the eigenvectors of the transition rate matrix to identify the steady-state distribution.

If the master equation (2.17) represents a thermodynamic system, the entries in the transition rate matrix are further restricted to satisfy the *local detailed balance* relationship

$$\frac{R_{ij}}{R_{ji}} = \exp(-\beta \Delta\omega_{ij}) \ , \tag{2.20}$$

where $\Delta\omega_{ij} = \omega_i - \omega_j$ is the difference in the thermodynamic potential of states x_i and x_j. For instance, if the system is in contact with a thermal reservoir, the thermodynamic potential of each state x_i is its energy $\omega_i = \epsilon_i$, and thus asymmetries in the transition rates are completely determined by the relative energies of the states. The local detailed balance constraint is necessary in such systems so that they remain consistent with the known predictions of thermodynamics. Specifically, for a thermodynamic system the equilibrium distribution is known from the study of statistical mechanics, and the rate definitions in (2.20) ensure that the stationary distribution and the equilibrium distribution are equal.

To generate trajectories of a system obeying the master equation (2.17), we need to know the *transition kernel* $p(x, t + \Delta t | x', t)$, which can be obtained by expanding the time-dependent probability distribution from (2.17) to linear order in Δt and using the master

equation (2.17) to evaluate the derivatives [72]:

$$p(x_i, t+\Delta t|x_j, t) = \delta_{ij}\left[1 - \sum_k R_{kj}\Delta t\right] + R_{ij}\Delta t + \mathcal{O}(\Delta t^2)\ , \tag{2.21}$$

where $\delta_{ij} = p(x_i, t|x_j, t)$. The first RHS term is the probability that during time Δt the system remains at x_j, while the second RHS term is the probability that it transitioned to state x_i from x_j. Thus, for a system initially in state x_j, after a small time Δt has passed, we can calculate its state by drawing a new state (x_i) at random from $p(x_i, t+\Delta t|x_j, t)$. Numerically, this can be achieved by implementing a kinetic Monte-Carlo algorithm.[4] The details of our implementation are given in App. A.

As an example, consider the underlying set of states $x_i \in \mathbb{Z}$, where $\mathbb{Z}$ is the set of integers, subject to a harmonic potential centered at $x_0 = 0$, so that that the system evolves in the potential $\epsilon_i = \frac{k_{\text{trap}}}{2}x_i^2$ where k_{trap} is the trap stiffness. Transition rates of the general form

$$R_{ij} = \Gamma e^{-\frac{1}{2}\beta\Delta\epsilon_{ij}} \tag{2.22}$$

for constant kinetic bare rate Γ and $\Delta\epsilon_{ij} \equiv \epsilon_i - \epsilon_j$ satisfy local detailed balance (2.20) and thus represent the stochastic dynamics of a physical system in contact with a thermal reservoir.

Figure 2.1 shows a sample trajectory (left) of the system in the harmonic potential for the first 50 discrete jumps, and the steady-state distribution p_x^{ss} (2.19) (right) given by the eigenvector corresponding to the zero-eigenvalue of the transition rate matrix. The trajectory consists of lag times, during which the system remains in a fixed state, interspersed with instantaneous discrete jumps.

2.2.2 Fokker-Planck equation

For systems in which there are no discrete jumps ($R(\boldsymbol{x}|\boldsymbol{x'}, t) = 0$ everywhere), but nonzero drift and diffusion coefficients, the differential Chapman-Kolmogorov equation (2.16) reduces to the *Fokker-Plank equation* which, in its general form, is

$$\partial_t p(\boldsymbol{x}, t) = -\sum_i \partial_{x^i}\left[A_i(\boldsymbol{x}, t)p(\boldsymbol{x}, t|\boldsymbol{x'}, t')\right] + \frac{1}{2}\sum_{i,j}\partial^2_{x^i,x^j}\left[B_{ij}(\boldsymbol{x}, t)p(\boldsymbol{x}, t|\boldsymbol{x'}, t')\right]\ . \tag{2.23}$$

For the duration of this thesis, we will be concerned with time- and space-independent diffusion, so that $B_{ij}(\boldsymbol{x}, t) = 2D$ (so that the factor of $1/2$ in front of the diffusive term is canceled), and further restrict our attention to the 1D Fokker-Planck equation corresponding

[4] This is analogous, in this case, to the Gillespie algorithm used to propagate the chemical master equation.

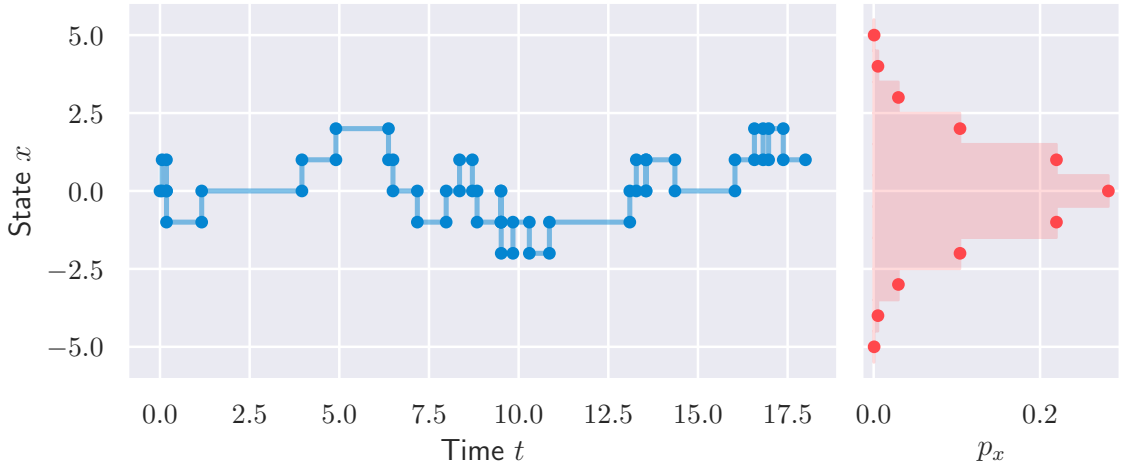

Figure 2.1: **Discrete-state dynamics generated by the master equation in a harmonic potential**. (left) Sample of the first 50 steps of a discrete-state trajectory governed by the discrete master equation (2.17) and subject to a harmonic confining potential. Each circle represents the start or end of a discrete jump in the state. We set $\Gamma = 1$ in the transition rates (2.22). (right) Empirical distribution (circles) obtained from a long trajectory of the discrete dynamics compared to the corresponding Boltzmann distribution (1.1) (bars) for a system in thermodynamic equilibrium. In both subplots we take $\beta = 1$.

to an overdamped system, which is also known as the Smoluchowski equation.[5] Here, we consider the Smoluchowski equation for a diffusing particle in a time-dependent potential $E(x,t)$, which means that the drift coefficient in (2.23) is replaced by (βD times) the force $f(x,t) \equiv -\partial_x E(x,t)$: $A(x,t) = -\beta D \partial_x E(x,t)$.

Integrating the Fokker-Planck equation over an (arbitrary) initial distribution $p(x, t = 0)$ converts the conditional equation into an equation for the marginal probability distribution $p(x,t)$. Thus the 1D overdamped Fokker-Planck equation is

$$\partial_t p(x,t) = \beta D \, \partial_x \left[f(x,t) p(x,t) \right] + D \, \partial_x^2 p(x,t) \; . \tag{2.24}$$

The stationary solution—if it exists—of the Fokker-Planck equation can be obtained by setting $\partial_t p(x,t) = 0$ and solving the RHS of (2.24).

In general, the full time-dependent distribution $p(x,t)$ is rarely analytically tractable, and must be obtained through numerical methods. However, when the energy landscape is harmonic ($E(x) = \frac{k_{\text{trap}}}{2} x^2$), the forces are linear, $f(x) = -\partial_x E(x) = -k_{\text{trap}} x$, and the

[5] To get an underdamped equation of motion, we need to also include the time evolution of the distribution of velocities. The Kramers equation does just this for a 1D system, but we won't discuss it in detail here. For the interested reader, however, an excellent overview is given in Hannes Risken's book [89].

Smoluchowski equation becomes

$$\partial_t p(x,t) = -\beta D k_{\text{trap}} \left[x p(x,t) \right] + D\, \partial_x^2 p(x,t) \ . \tag{2.25}$$

Here—and, in fact, for any Fokker-Planck equation subject to a potential-derived drift coefficient—the (equilibrium) steady-state solution can be found from the Boltzmann distribution

$$p_{\text{eq}}(x) = \sqrt{\frac{k_{\text{trap}}}{2\pi}} e^{-\frac{\beta k_{\text{trap}}}{2} x^2} \ , \tag{2.26}$$

which here is Gaussian with mean $\langle x \rangle_{\text{eq}} = 0$ and variance $\langle \delta x^2 \rangle_{\text{eq}} = (\beta k_{\text{trap}})^{-1}$. Here, the angle brackets $\langle \cdots \rangle_{\text{eq}}$ indicate an average over the equilibrium distribution (2.26).

In this special case we can also find an analytical solution for the full time-dependent distribution, for any initial conditions. For an initial distribution that is localized at $x = x'$ so that $p(x, t = 0) = \delta(x - x')$, the time-dependent solution is Gaussian at all times $t > 0$,

$$p(x,t|x',t'=0) = \sqrt{\frac{\beta k_{\text{trap}}}{2\pi(1 - e^{-2\beta D k_{\text{trap}} t})}} \exp\left[-\frac{\beta k_{\text{trap}}}{2} \frac{(x - x' e^{-\beta D k_{\text{trap}} t})^2}{1 - e^{-2\beta D k_{\text{trap}} t}} \right] \ , \tag{2.27}$$

with mean $\langle x(t) \rangle = x' e^{-\beta D k_{\text{trap}} t}$ and variance $\langle \delta x^2 \rangle = (1 - e^{-2\beta D k_{\text{trap}} t})/(\beta k_{\text{trap}})$. Asymptotically as $t \to \infty$, the mean and variance approach the steady-state values. We will make extensive use of this general form in Chapter 7.

Figure 2.2 shows the time-dependent probability distribution obtained by solving the Smoluchowski equation (2.24) in a harmonic potential with $\beta k_{\text{trap}} = D = 1$, given an initial distribution $p(x, t = 0) = \delta(x - 2)$. The distribution relaxes towards the equilibrium distribution—indicated by a dashed red line—obtained from the Boltzmann equation (2.26). The right panel shows the time-dependent average and variance of the distribution, which converge towards their equilibrium values of $\langle x \rangle_{\text{eq}} = 0$ and $\langle \delta x^2 \rangle_{\text{eq}} = 1$.

2.2.3 Langevin equation

Similar to the use of the transition kernel of the master equation to generate trajectories that represent the time evolution of single instances of the stochastic process, the transition kernel of the Fokker-Planck equation gives rise to a stochastic equation of motion for individual trajectories known as the Langevin equation. For instance, one can show that for the Smoluchowski equation of a diffusing particle in a potential $E(x,t)$, the transition kernel is, to leading order in Δt [72],

$$p(x, t + \Delta t|x', t) = \sqrt{\frac{D}{2\pi \Delta t}} \exp\left[-\frac{(x - x' + \beta D f(x', t) \Delta t)^2}{2 D \Delta t} \right] \ , \tag{2.28}$$

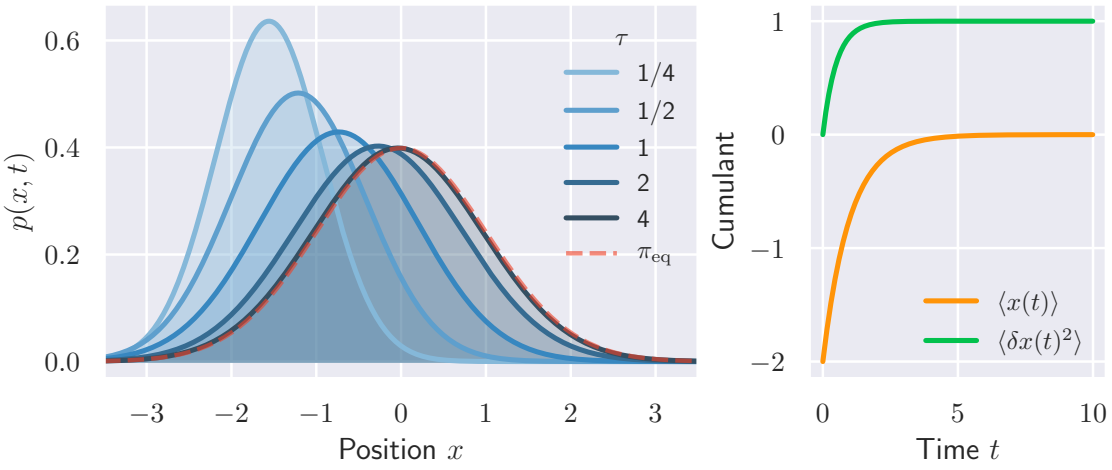

Figure 2.2: **Time-dependent solution to the Smoluchowski equation for a diffusing particle in a harmonic potential.** (left) Time-dependent solution to the Smoluchowski equation for a Brownian particle diffusing in a harmonic potential (2.27), given an initial distribution $p(x,0) = \delta(x-2)$. Early times are indicated by light blue, while later times are indicated by progressively darker shades of blue. The red dashed line is the Boltzmann equilibrium distribution towards which the system is relaxing. (right) The values of the first two cumulants—mean (yellow) and variance (green)—of the distribution $p(x,t)$ as a function of time, showing that they asymptote to their equilibrium values of $\langle x\rangle = 0$ and $\langle \delta x^2\rangle = 1$ as $t \to \infty$.

which is simply a Gaussian distribution with variance $D\Delta t$ and mean $x' - \beta D f(x',t)\Delta t$, where $f(x',t) = -\partial_x E(x,t)$ is the force experienced by the particle at position x and time t. Thus, the resulting stochastic dynamics are those of a system moving with a systematic drift of velocity $\beta D f(x',t)$ with zero-mean Gaussian fluctuations with variance $D\Delta t$.

Therefore, the update rule for a trajectory generated by this equation of motion is

$$x(t+\Delta t) = x(t) + \beta D f(x,t)\Delta t + \sqrt{2D}\mathcal{N}(0,\Delta t)\ , \tag{2.29}$$

where $\mathcal{N}(0,\Delta t)$ is a normal distribution with 0 mean and variance Δt. The Gaussian-fluctuation term is known as a *Wiener increment*, and often indicated by ΔW. In the continuous-time limit ($\Delta t \to 0$), (2.29) becomes

$$\mathrm{d}x = \beta D f(x,t)\mathrm{d}t + \sqrt{2D}\,\mathrm{d}W(t)\ , \tag{2.30}$$

where $\mathrm{d}W(t)$ is the differential element of a continuous-time stochastic process $W(t)$, known as a *Wiener process* (also called a white-noise process as it has a flat power spectral density over all frequencies) [72]. The Wiener process is defined by a couple of important properties, namely it has zero mean ($\langle W(t)\rangle = 0$), unit variance ($\langle W(t)^2\rangle = 1$), and its values at different times are independent, $\langle W(t)W(t')\rangle = \delta(t-t')$.

Alternatively, it is common to rewrite (2.30) as a differential equation, in terms of derivatives as opposed to differentials:

$$\dot{x} = \beta D f(x,t) + \sqrt{2D}\,\xi(t) \ , \tag{2.31}$$

where $\dot{x} \equiv \mathrm{d}x/\mathrm{d}t$, $\xi(t) \equiv \mathrm{d}W(t)/\mathrm{d}t$, and $\langle \delta\xi(t)\delta\xi(t')\rangle = \delta(t-t')$. Throughout this thesis, we will primarily use the derivative form (2.31), except when using the differential form is more clear.

Equation (2.30) (or (2.31)) is an example of a stochastic differential equation (SDE) because it contains a stochastic differential term ($\mathrm{d}W(t)$ or $\xi(t)$). As we will see, this makes the subsequent analysis significantly different from standard ordinary differential equations. Here, the sample paths generated by the Langevin equation (2.30) are continuous everywhere, but differentiable nowhere, due to the properties of the Wiener process [72].[6]

The trajectories generated by the SDE in (2.30) represent the same underlying physical process as the Smoluchowski equation (2.24). In both cases we are implicitly assuming that the system is *overdamped*. This means that there are no inertial terms in the equations of motion, which is an oft-used approximation in low-Reynolds number environments.

The Reynolds number 'Re' is a dimensionless quantity used in fluid mechanics to classify the relative importance of inertial and viscous forces [90]. For small values of the Reynolds number, viscous forces dominate, and inertial effects can largely be ignored [91]. Approximations on a single bacterium predict that its Reynolds number is $\sim 10^{-5}$ which implies that its dynamics are governed largely by viscous forces.[7] Thus, the overdamped approximation plays a prominent role in the microscopic physics of biological systems.

The overdamped equation of motion (2.30) is a particular limit of a more general Langevin equation, where the fluctuations, which are assumed to impart forces on the microscopic particle, appear in the equation of motion for the acceleration. Specifically, the underdamped Langevin equation, which could be derived, in principle, from the transition kernel for a Fokker-Planck equation of an underdamped system–such as the Kramers equation–is

$$m\ddot{x} = -f(x,t) - \frac{1}{\beta D}\dot{x} + \sqrt{\frac{2}{\beta^2 D}}\xi(t) \ . \tag{2.32}$$

[6] In fact, this is an argument made by Jacobs in [81] for the use of differential notation for SDEs in preference to writing them in terms of derivatives. The non-differentiability of sample paths implies that the derivatives themselves don't exist, and thus it is more mathematically consistent to write the Langevin equation (or any SDE for that matter) in differential notation, such as in (2.30).

[7] Specifically, the Reynolds number is defined mathematically as $\mathrm{Re} \equiv Ua/\nu$ where U is the speed of the object, a is a characteristic linear dimension, and ν is the kinematic viscosity of the fluid (for water, $\nu \approx 10^{-6}$ m^2/s). For a typical bacterium (such as, for instance, *E. coli*) the characteristic linear dimension is $a \approx 10^{-6}$ m and they travel at velocities of $U \approx 10^{-5}$ m/s and thus the Reynolds number is $\mathrm{Re} \approx 10^{-5}$ [91].

Here, the first term on the RHS represents the force applied to the system (either by a potential, or some external source), the second term on the RHS represents the frictional damping of velocities, and the final term on the RHS represents the stochastic forces on the particle from the environment. To recover the overdamped equation of motion (2.31) from the underdamped equation of motion, we simply set $\ddot{x} = 0$ and solve for $\dot{x}$.

2.3 Stochastic thermodynamics

The preceding sections describe methods to generate the stochastic dynamics of fluctuating systems, based solely on the mathematical concepts of probability theory; however, these trajectories still lack a thermodynamic interpretation. This interpretation is provided by the growing field of *stochastic thermodynamics*. Stochastic thermodynamics is the study of thermodynamics at the microscopic scale, where systems are small and fluctuations cannot be ignored. At its heart, stochastic thermodynamics seeks to generalize the classical laws of thermodynamics to fluctuating trajectories. As considered here, the systems of interest are embedded within an aqueous environment, and can be pushed out of equilibrium in a number of different ways. For instance, systems can be driven out of equilibrium by altering their potential energy in a time-dependent manner through a *control parameter* λ, which we discuss in more detail in Sec. 2.6. Here, we can study the response of the system to a time-dependent control parameter $\lambda(t)$, the spontaneous relaxation towards equilibrium after an instantaneous change in the control parameter value, or the properties of nonequilibrium steady states when the driving is time-independent.

Alternatively, systems such as chemical reaction networks can be kept out of equilibrium by chemostats that maintain the out-of-equilibrium concentrations of chemical reactants. Here, the system is still driven out of equilibrium, even though there is no active manipulation of the system through, for instance, a control parameter.

Formally, the separation of system from surroundings in stochastic thermodynamics involves identifying fast- and slow-relaxing degrees of freedom. The slow-relaxing degrees of freedom constitute the system, while the fast-relaxing degrees of freedom comprise the bath (or environment). This is one reason why the study of microscopic biological systems has proved to be difficult: there is often no obvious separation of timescales allowing one to separate, unambiguously, the system from the environment.

At the microscopic scale, the mathematical forms of the laws of thermodynamics are largely preserved, but a complete understanding of how exactly to attribute increments of work, heat, and entropy to single trajectories is a nontrivial process. Throughout this thesis we will adopt the convention that positive work and heat correspond to energy flows into the system. Naturally, this gives rise to the differential form of the first law

$$\mathrm{d}E = đW + đQ \ , \tag{2.33}$$

where $\mathrm{d}E$ represents a change in the internal energy of the system, and the inexact differentials đ indicate that the work and heat are path-dependent.

Given an overdamped Langevin equation describing a fluctuating particle, Sekimoto was the first to suggest a thermodynamic interpretation [92]. In particular, how can we quantify the heat and work associated with a single stochastic trajectory? Conceptually, work is done on a system when an external agent inputs energy, and heat represents the exchange of energy between the system and its environment. Thus, in the absence of any external input, all energy changes are heat. For an overdamped system fluctuating in a fixed potential energy landscape $V(x, \lambda)$, changes in energy are determined by the changes in potential $\mathrm{d}E = \mathrm{d}V$, which, by using the chain rule, can be written in terms of a differential in the particle position $\mathrm{d}V = (\partial_x V)\mathrm{d}x = -đQ$, quantifying the exchange of heat with the reservoir. Thus, the heat along a stochastic trajectory is given by the integral expression [27, 92]

$$Q[x_{[0,\tau]}] = -\int_0^\tau f(x(t))\dot{x}\,\mathrm{d}t\ , \tag{2.34}$$

where here the notation $Q[x_{[0,\tau]}]$ indicates that the heat is a functional of the trajectory x during times $t \in [0, \tau]$. We can use (2.34) to obtain an exact expression for the heat in terms of the stochastic dynamics by substituting (2.31) for $\dot{x}$. However, because $\dot{x}$ is a stochastic differential equation, direct integration could pose a problem, as the integral of the stochastic process has not yet been defined. We omit discussion of the different definitions of stochastic integrals, but specify here that the Sekimoto definitions of heat (2.34) (and work (2.35)) must use the Stratonovich interpretation of the stochastic integral [92].

Alternatively, work is done on the system if the control parameter λ changes in time. Thus, changes in energy due to work are $\mathrm{d}E = \mathrm{d}V = (\partial_\lambda V)\mathrm{d}\lambda = đW$, and the work done along a trajectory is

$$W[x_{[0,\tau]}] = \int_0^\tau (\partial_\lambda V(x(t)))\,\dot{\lambda}\,\mathrm{d}t\ . \tag{2.35}$$

Combining (2.35) with (2.34), the integrated first law is

$$\Delta E = V(x(\tau), \tau) - V(x(0), 0) = Q[x_{[0,\tau]}] + W[x_{[0,\tau]}]\ , \tag{2.36}$$

on the level of an individual stochastic trajectory of the system.

Equation (2.36) provides a powerful conceptual separation of the work and heat accumulation at the microscopic level: work accumulates when external influences on the system results in changes of energy, while heat accumulates when the system exchanges energy with the bath. The first law, which effectively is a statement of energy conservation, links these two energy flows.

Identifying a valid form of the second law of thermodynamics requires a consistent definition of entropy along a single trajectory. However, the second law refers to the entropy of the universe, so we must identify the changes in entropy of both the system and the

environment. The entropy change in the environment during a trajectory can be identified with the heat

$$\Delta S_{\rm res}[x_{0,\tau}] = -\beta Q[x_{[0,\tau]}], \tag{2.37}$$

while the instantaneous entropy of the system at time t is

$$S(t) = -\ln p(x(t), t) \ , \tag{2.38}$$

where, here and throughout the thesis, we use natural units ($k_{\rm B} = 1$) for the entropy. Here, the probability distribution $p(x(t), t)$ is obtained by first solving, for instance, the corresponding Fokker-Planck equation for the system dynamics. The second law thus follows by measuring the total change in entropy of the universe over a trajectory,

$$\Delta S_{\rm tot}[x_{[0,\tau]}] = \Delta S_{\rm res}[x_{[0,\tau]}] + S(\tau) - S(0) \ , \tag{2.39}$$

where, the final two terms on the RHS capture the change in the entropy of the system state over the time interval. Unlike the usual form of the second law, the entropy change along a single trajectory can be negative [93]. For consistency with the classical laws of thermodynamics, we require only that $\langle \Delta S_{\rm tot}[x_{0,\tau}] \rangle_\Lambda \geq 0$ where $\langle \cdots \rangle_\Lambda$ indicates an average over an ensemble of trajectories under the same driving. In (2.39), there is no requirement that the system starts and/or ends in equilibrium, however to calculate the entropy from (2.38) we must know the probability distribution $p(x(t), t)$ over states.

The stochastic thermodynamic study of trajectory-level entropy led to a far more general form of the second law through an entirely new class of mathematical identities known collectively as *fluctuation theorems*.

2.4 Fluctuation theorems

Historically, the identification of symmetries under time reversal of dynamical systems began with a seminal paper by Gallavotti and Cohen in which they related the change in Kolmogorov-Sinai entropy[8] to the ratio of trajectory probabilities under time-forward and time-reversed dynamics [8]. Subsequent study on the probabilities of second-law-violating events in steady-state systems [9, 10] as well as earlier research on the work integrated over driven trajectories by Bochkov and Kuzovlev [95, 96] led to the well-known Jarzynski equality [6],

$$\langle e^{-\beta W} \rangle_\Lambda = e^{-\beta \Delta F} \ , \tag{2.40}$$

[8]The Kolmogorov-Sinai entropy is a definition of entropy, used often in dynamical systems theory, that is calculated from the Lyapunov exponents in chaotic systems [94].

which relates the exponentiated trajectory work $\beta W \equiv \beta W[x_{[0\tau]}|\Lambda]$ averaged over an ensemble of nonequilibrium trajectories (LHS), to the exponentiated change in equilibrium free energy $\beta\Delta F \equiv \beta F(\tau) - \beta F(0)$ (RHS). In the Jarzynski equality, the system is assumed to be driven through the manipulation of an external control parameter λ—or, in Jarzynski's words, a *work parameter*—and the average $\langle\cdots\rangle_\Lambda$ is taken over the ensemble of nonequilibrium responses of the system to the particular driving protocol. Importantly, the Jarzynski equality holds regardless of how far the system is driven out of thermodynamic equilibrium.

By making an appeal to Jensen's inequality—an inequality pertaining to averages over convex functions that is used ubiquitously in information theory [94]—the Jarzynski equality can be reduced to the inequality,

$$\langle W\rangle \geq \Delta F \tag{2.41}$$

or, alternatively

$$\langle W_{\text{ex}}\rangle \geq 0 \tag{2.42}$$

where here we define the *excess work* $W_{\text{ex}} \equiv W - \Delta F$ as the difference between the work and the free energy difference. The lower bounds implied by the Jarzynski equality (2.41) (and (2.42)) can be seen as alternative statements of the second law of thermodynamics, when the excess work is the only source of entropy production [93].

The Jarzynski equality is an example of what is now called an *integrated fluctuation theorem* (IFT), to contrast with a *detailed fluctuation theorem* (DFT). While an IFT involves ensemble averages, a DFT relates the probabilities of individual stochastic trajectories. The first DFT found was the Crooks fluctuation theorem [7],

$$\frac{p[x|\Lambda]}{\tilde{p}[\tilde{x}|\tilde{\Lambda}]} = e^{-\beta W_{\text{ex}}[x|\Lambda]}\ , \tag{2.43}$$

which relates the ratio of trajectory probabilities $p[x|\Lambda]$ for a system driven through a time-dependent control protocol Λ, to the probability of its time-reversed trajectory $\tilde{p}[\tilde{x}|\tilde{\Lambda}]$ under the time-reversed protocol $\tilde{\Lambda}$. Here the tilde indicates the time-reversal operation, with $\tilde{x}(t) = x(\tau - t)$ and $\tilde{\lambda}(t) \equiv \lambda(\tau - t)$. The Crooks theorem can be re-written in terms of the work distribution along forward and reverse processes as $p(W_{\text{ex}})/\tilde{p}(-W_{\text{ex}}) = \exp(\beta W_{\text{ex}})$ and, in fact, can be used to derive the Jarzynski equality. Thus, the Crooks theorem represents the DFT that, upon integration, reproduces the Jarzynski IFT. In fact, the existence of a DFT directly implies a corresponding IFT, and vice-versa [97].

Among the fluctuation theorems, the entropy production fluctuation theorem is central, reducing to other fluctuation theorems, such as the Crooks IFT, in particular settings. Proven in its current form by Seifert in 2005 [12], the entropy production fluctuation theorem is

$$\frac{p(\Delta S_{\text{tot}})}{\tilde{p}(-\Delta S_{\text{tot}})} = e^{\Delta S_{\text{tot}}}\ , \tag{2.44}$$

relating the exponentiated change in entropy to the ratio of the probability of observing a total entropy change ΔS_{tot} along a forward trajectory to the probability of observing a corresponding decrease in entropy under the time-reversed process. For a system driven by an external control parameter, the excess work (nondimensionalized by β) and entropy production are equal, thus (2.44) reduces to the Crooks fluctuation theorem (2.43). In fact many of the well-known fluctuation theorems can be recovered from (2.44) by splitting the change in entropy into *adiabatic* and *nonadiabatic* contributions, $\Delta S_{\text{tot}} = \Delta S_{\text{a}} + \Delta S_{\text{na}}$, capturing the dissipation due to external time-dependent driving ΔS_{na} (such as by a time-dependent control parameter) or due to nonequilibrium boundary conditions ΔS_{a} (such as out-of-equilibrium concentrations of chemical reactants), respectively [97, 98, 99].

Furthermore, by using Jensen's inequality, (2.44) reduces to the familiar form of the second law of thermodynamics, $\langle \Delta S_{\text{tot}} \rangle \geq 0$. Thus, (2.44) represents a true generalization of the second law of thermodynamics, which only became apparent through the study of microscopic fluctuating systems. The exponential suppression of 'second-law violating' trajectories makes the probability of observing such events in macroscopic systems vanishingly small [93]. Broadly speaking, the fluctuation theorems have provided a far deeper understanding of the microscopic physics of fluctuating systems, placing surprisingly general constraints on the fluctuations and dissipation in thermodynamic systems, even far from equilibrium.

2.5 Entropy and information theory

From its inception in the 1940s by Claude Shannon at Bell Labs, the field of information theory provided a comprehensive and powerful set of tools to describe the mathematics of information. More recently, the field of nonequilibrium statistical mechanics and stochastic systems have found several concepts from information theory essential for understanding the physics of nonequilibrium systems. Here, we briefly review some of the relevant results from information theory.

The idea of entropy is central to information theory where, unlike many classical treatments in thermodynamics, it possesses a conceptually simple explanation. The entropy is a measure of the amount of information required, on average, to describe the state of a system. Low-entropy states (such as a long sequence of heads in successive coin flips) are highly ordered, and thus require a relatively small amount of information to specify, whereas high-entropy states are more disordered, requiring more information to completely specify.[9] Mathematically, the entropy of a random variable X—or Shannon entropy as it is typically

[9]This explanation of the entropy is clarified when the entropy is described in bits, by using base-2 logarithms. Here, the entropy is literally the average number of bits—yes or no questions—required to completely specify the state of a system.

referred to—is

$$S(X) \equiv -\sum_x p_x \ln p_x \ . \tag{2.45}$$

When multiple random variables (X and Y) are involved, one can also define the joint and conditional entropy respectively as

$$S(X,Y) = -\sum_{x,y} p_{xy} \ln p_{xy} \tag{2.46a}$$

$$S(X|Y) = -\sum_y p_y \sum_x p_{x|y} \ln p_{x|y} \ . \tag{2.46b}$$

The conditional entropy (2.46b) is useful in decomposing the joint entropy (2.46a) via the chain rule for entropies [94]:

$$S(X,Y) = S(X) + S(Y|X) \tag{2.47a}$$

$$= S(Y) + S(X|Y) \ . \tag{2.47b}$$

Central to the study of information theory is the measure of *mutual information*,

$$I(X;Y) \equiv \sum_{x,y} p_{xy} \ln \frac{p_{xy}}{p_x p_y} \ , \tag{2.48}$$

which quantifies the information that is shared between two random variables. The mutual information (2.48) is symmetric ($I(X;Y) = I(Y;X)$), nonnegative, and can be used to relate the marginal entropy of a single random variable in a multivariable system to its conditional entropy:

$$I(X;Y) = S(X) - S(X|Y) \ . \tag{2.49}$$

This provides a conceptual interpretation of $I(X;Y)$ as the reduction in uncertainty (entropy) about a random variable X by observation of another random variable Y. If X and Y are independent, then $p_{xy} = p_x p_y$ and $I(X;Y) = 0$ (2.48). Conversely, if X and Y are correlated, then the non-negativity of mutual information requires that $S(X|Y) < S(X)$.

Information theory also provides a number of so-called divergence measures that quantify the differences between probability distributions. Such divergences have become central to the study of nonequilibrium systems, in particular for lower bounding statistical averages. The *relative entropy* (or *Kullback-Leibler* divergence) is a common measure used in nonequilibrium statistical physics, and is defined for probability distributions p_x and q_x as

$$D\left(p_x||q_x\right) \equiv \sum_x p_x \ln \frac{p_x}{q_x} \tag{2.50}$$

The relative entropy is nonnegative for any two probability distributions, $D(p_x||q_x) \geq 0$—a property which is commonly used to prove the non-negativity of physical quantities in nonequilibrium thermodynamics—but is asymmetric in its arguments ($D(p_x||q_x) \neq D(q_x||p_x)$ in general) [94]. However, one can *symmetrize* the relative entropy in several different ways, such as taking the relative entropy of each distribution to their arithmetic mean $m_x \equiv \frac{1}{2}(p_x + q_x)$ to get the Jensen-Shannon divergence [100]

$$\mathrm{JSD}\left(p_x, q_x\right) \equiv \frac{1}{2} D\left(p_x||m_x\right) + \frac{1}{2} D\left(q_x||m_x\right) \ . \tag{2.51}$$

The divergence measures originally derived in the context of information theory arise, in many forms, within the study of nonequilibrium systems. For instance, we will see in Chapter 7 how the relative entropy is related to excess work in discretely driven systems. In fact, the physical meanings of various divergence measures in the context of trajectory ensembles in stochastic thermodynamics have been explored in detail in [100], relating the relative entropy (2.50) to dissipation, Jensen-Shannon divergence (2.51) to time asymmetry, and providing physical interpretations of various other divergence measures from information theory in the context of thermodynamic trajectories.

When a system is controlled by a set of external parameters $\boldsymbol{\lambda}$, the equilibrium distribution $\pi_x(\boldsymbol{\lambda})$ is parameterized by the control parameters $\boldsymbol{\lambda}$ (here, assumed to take on any of a continuous range of values). Thus, we can define a quantitative measure of the difference between two equilibrium distributions by taking the relative entropy $D[\pi_x(\boldsymbol{\lambda})||\pi_x(\boldsymbol{\lambda}')]$. If the two distributions are close to one another, so that $\boldsymbol{\lambda}' \approx \boldsymbol{\lambda} + \Delta\boldsymbol{\lambda}$, then the Taylor expansion of the relative entropy is, to leading order in $\Delta\boldsymbol{\lambda}$,

$$D[\pi_x(\boldsymbol{\lambda})||\pi_x(\boldsymbol{\lambda}')] \approx \frac{1}{2} \Delta\lambda^i \Delta\lambda^j \, \mathcal{I}_{ij}(\boldsymbol{\lambda}) \ . \tag{2.52}$$

Here,

$$\mathrm{d}I_{ij}(\boldsymbol{\lambda}) \equiv \langle \partial_{\lambda^i} \ln \pi_x(\boldsymbol{\lambda}) \, \partial_{\lambda^j} \ln \pi_x(\boldsymbol{\lambda}) \rangle_{\boldsymbol{\lambda}} \tag{2.53}$$

is the (i,j)th component of the *Fisher information matrix*, and $\Delta\lambda^i$ is the ith component of $\Delta\boldsymbol{\lambda}$ [94]. Here, and throughout the remainder of the thesis we have employed an Einstein summation notation, where repeated indices are implicitly summed over, for instance, $\Delta\lambda^i \Delta\lambda^j \mathcal{I}_{ij}(\boldsymbol{\lambda}) \equiv \sum_{i,j} \Delta\lambda^i \Delta\lambda^j \mathcal{I}_{ij}(\boldsymbol{\lambda})$. The Fisher information plays an important role in quantifying the dissipation in driven, nonequilibrium processes, and will be encountered periodically throughout the remainder of this thesis.

2.6 Control in microscopic nonequilibrium systems

Broadly speaking, control theory is the study of physical systems and how their state can be controlled through interaction with the outside world. By identifying particular goals,

it is then possible to design ways to exercise control over a system to achieve them [101]. For instance, a canonical example of a mechanical controller is the *centrifugal governor*—an idea which was explored in detail by James Clerk Maxwell in the 1800s [102]—where the motion of a centrifugal rotor, coupled to the engine output, regulates engine speed by controlling the input valve. Since its early applications, the scope of control theory has expanded greatly and plays a major role in the modern practice of science, particularly in the context of experimental physics, where many techniques rely fundamentally on the successful implementation of control-theory-based feedback techniques [103].

Optimal control—one subfield of the broader subject of control theory—deals with the implementation of control strategies that minimize some measure of cost in a controlled system. Through much of this thesis, we will deal with the utilization of optimal control to minimize the amount of work required to manipulate stochastic systems in a particular way.[10] In particular, we will often consider stochastic systems which fluctuate in the presence of a potential energy $E(x, \boldsymbol{\lambda})$ which is determined by a set of control parameters $\boldsymbol{\lambda}$. The instantaneous value of $\boldsymbol{\lambda}$ determines the equilibrium state $\pi(x|\boldsymbol{\lambda})$ of the system through (1.1). A *control protocol* $\Lambda \equiv \boldsymbol{\lambda}_{[0,\tau]}$ is a time-dependent schedule for the control parameter $\boldsymbol{\lambda}$, driving the system from an initial state—often the equilibrium state at $\boldsymbol{\lambda}(t = 0)$—to a final state at $\boldsymbol{\lambda}(t = \tau)$, where the protocol duration is denoted by τ.

Following from the Sekimoto definition (2.35), work is done on the system (or extracted from the system) whenever the control parameter is changed. For a particular time-dependent control protocol Λ that transforms $\boldsymbol{\lambda}_\mathrm{i} \to \boldsymbol{\lambda}_\mathrm{f}$ over a protocol duration τ, the mean work required is given by the functional

$$\langle W \rangle_\Lambda = -\int_0^\tau \langle f_i(x, \boldsymbol{\lambda}) \rangle_{\Lambda(t)} \dot{\lambda}^i \, \mathrm{d}t \ , \tag{2.54}$$

where $f_i \equiv -\partial_{\lambda^i} E(x, \boldsymbol{\lambda})$ is the generalized force conjugate to control parameter λ^i, the angle brackets $\langle \cdots \rangle_{\Lambda(t)}$ indicate an average over the instantaneous (nonequilibrium) distribution of the system at time t during the protocol Λ, and the angle brackets $\langle \cdots \rangle_\Lambda$ indicate an average over the (nonequilibrium) distribution of system trajectories during the entire control protocol Λ. Given that the initial and final control parameter values are generally fixed,[11] the equilibrium free energy difference $\Delta F_\Lambda \equiv F(\boldsymbol{\lambda}_\mathrm{f}) - F(\boldsymbol{\lambda}_\mathrm{i})$ is constant for all protocols, and thus minimizing the work is the same as minimizing the excess work $W_\mathrm{ex} = W - \Delta F$. For the remainder of this thesis, we will focus on minimizing the excess work.

[10] Actually, this represents a further subfield known as *stochastic optimal control*, as it pertains to optimal control strategies in stochastic systems. However, in Chapter 6 we will refer to the control of systems through stochastic driving protocols as 'stochastic control', and thus, to avoid confusion, we will refer to stochastic optimal control in this context simply as optimal control.

[11] Recent work on optimal bit erasure [104, 105] enforces a constraint on the final probability distribution rather than requiring that the control parameter reaches a particular final value.

For a system bound by a harmonic potential (see Sec. 2.7.1), Schmiedl and Seifert [106] found the exact optimal control protocols that minimize the average work when the control parameter is either the position of the harmonic potential, $E(x,\lambda) = \frac{k_{\text{trap}}}{2}(x-\lambda)^2$, or the trap strength, $E(x,\lambda) = \frac{\lambda}{2}x^2$ [106]. Further, in 2011, Aurell, *et al.* expanded the set of exactly solvable minimum-work protocols to a more general class of problems by finding a clever mapping of the stochastic differential equations to a set of deterministic transport equations [107, 108]. While this work provided a new means of analyzing such problems, aside from some notable recent results regarding a finite-time Landauer limit [104, 105], the study of analytically tractable solutions in optimal control of stochastic systems has seen little progress since. Primarily, this is due to the intrinsic difficulty in solving the general minimization problem in (2.54).

2.6.1 Linear-response theory

Aside from certain simple models, the search for exact solutions to the problem of optimal control in stochastic systems has remained relatively stagnant over the past half-decade. However, in the interim, phenomenological methods of approximating the average excess work associated with a particular driving protocol Λ have seen significant success, boasting predictions that are straightforward, simple to interpret, and easily scalable [109]. Often, the physical framework which underlies these approximations is linear-response theory.

Linear-response theory is a phenomenological approach to studying the near-equilibrium behaviour of systems responding to an external perturbation. In the context of stochastic thermodynamics, linear-response theory provides a mathematical machinery to approximate the nonequilibrium average of a physical quantity with equilibrium averages. This can be useful, for instance, when approximating the average work done on a system (2.54).

There are two classes of linear-response that—although closely related—are typically approached differently: linear-response for step perturbations and linear-response for time-dependent protocols. The former is primarily concerned with how an equilibrium system responds to a discrete external perturbation, while the latter estimates the response of a system to a continuous, time-dependent history of perturbations. Given the repeated use of both linear-response approaches throughout this thesis, we present a rather in-depth treatment of both here, and elaborate where necessary through the remainder of the thesis.[12]

Linear-response for step perturbations

Consider a thermodynamic system with dynamics governed by the potential energy $E_0(x)$. In the distant past ($t \ll 0$), the system was exposed to a weak external scalar field ϕ so

[12] David Chandler gives a particularly clear exposition on linear-response for step perturbations for Hamiltonian systems in [110], while Robert Zwanzig gives a clear overview of linear-response for time-dependent protocols in [5].

that, up to time $t = 0$, the effective potential energy has been $E(x, \phi) = E_0(x) + \phi f(x)$, where $f(x) \equiv -\partial_\phi E(x, \phi)|_{\phi=0}$ is the generalized force conjugate to the perturbing field ϕ.[13] (A canonical example of static linear response is the time-dependent response of an electric dipole moment in the presence of an applied electric field [5].) The system completely equilibrates with the effective potential energy $E(x, \phi)$, so that at time $t = 0$ its distribution over microstates is the Boltzmann distribution (1.1) for $E(x, \phi)$. For a general observable $A(x)$, the ensemble average at $t = 0$ is thus

$$\langle A \rangle_{\text{eq}} = \frac{\int A(x) e^{-\beta E_0(x) - \beta \phi f(x)} \mathrm{d}x}{\int e^{-\beta E(x) - \beta \phi f(x)} \mathrm{d}x} \ . \tag{2.55}$$

At time $t = 0$ the external field is instantaneously turned off, so that the dynamics are now governed by $E_0(x)$. We then want to approximate the relaxation of the observable $A(x)$ towards its new equilibrium average in the absence of the perturbing field ϕ.

For times $t \geq 0$, one can show that—for sufficiently weak perturbing fields ϕ—the time-dependent response of $A(x)$, averaged over the nonequilibrium distribution is (to linear order in $\mathcal{O}(\phi f(x))$) [110]

$$\langle A(t) \rangle_{\text{neq}} \approx \langle A \rangle_{\text{eq}} - \beta \phi \langle \delta A(t) \delta f(0) \rangle_{\text{eq}} \tag{2.56}$$

where the angle brackets $\langle \cdots \rangle_{\text{eq}}$ indicate an equilibrium average, and $\delta A(t) \equiv A(t) - \langle A \rangle_{\text{eq}}$ indicates a fluctuation of the observable A from its equilibrium value $\langle A \rangle_{\text{eq}}$.[14] Equation (2.56) shows that, in the linear-response regime, it is possible to approximate the nonequilibrium relaxation of an arbitrary observable $A(x)$ as a function of time t, given only observations of the equilibrium dynamics. Furthermore, we are often interested in the fluctuation of $A(x)$ away from its equilibrium value, and thus by subtracting $\langle A \rangle_{\text{eq}}$ from both sides of (2.56) we find

$$\langle \delta A(t) \rangle_{\text{neq}} \approx -\beta \phi \langle \delta A(t) \delta f(0) \rangle_{\text{eq}} \ . \tag{2.57}$$

This result is used extensively in the work presented in Chapter 7, where we also show how the static linear-response result (2.57) is related to dynamic linear response.

[13] Put another way, the weakness of the perturbing field ϕ ensures that it only couples linearly to the conjugate force.

[14] Originally, this result was derived for systems with Hamiltonian dynamics (as in, for instance, Ref. [110]), which led to skepticism about its broad applicability in microscopic systems (including a well-known critique from van Kampen in [111]). However, the linear-response predictions seemed to remain consistent with experimental observations in far more general situations than the Hamiltonian derivation would suggest. More recently, the analogous result has been derived for stochastic dynamics and nonequilibrium systems, however it is still an active area of research [112].

Furthermore, although we have assumed that the field ϕ and observable A are scalar quantities, the results presented above can be applied equally well to multidimensional fields and observables, with the nonequilibrium average of the ith component of observable $\boldsymbol{A}(x)$ relaxing from the jth perturbing field ϕ^j is

$$\langle \delta A_i(t) \rangle_{\text{neq}} \approx -\beta \phi^j \langle \delta f_j(0) \delta A_i(t) \rangle_{\text{eq}} \ , \tag{2.58}$$

where f_j is the generalized force conjugate to the jth component of the vector $\boldsymbol{\phi}$ [110].

Linear-response for time-dependent protocols

To now generalize the linear-response results for a step perturbation to a time-dependent protocol $\boldsymbol{\phi}(t)$, we can proceed in several different ways. In particular, one can expand the Liouville equation of the system's temporal evolution in phase space as in [5], or treat the system in a control-theoretical sense, resolving the relation between field (input) and observable (output) by expanding the response in a Volterra series and truncating to linear order [113]. In any case, the linear-response approximation of the nonequilibrium average is given by the integral expression

$$\langle \delta A_i(t) \rangle_{\text{neq}} \approx \int_{-\infty}^{0} \chi_{ij}(t - t') \left[\phi^j(t') - \phi^j(t) \right] \, \mathrm{d}t' \ , \tag{2.59}$$

where $\langle \delta A_i(t) \rangle_{\text{neq}}$ represents a fluctuation of A_i at time t away from its equilibrium average at $\phi(t)$, and $\chi_{ij}(t - t') \equiv \beta \frac{d}{dt} \langle \delta f_j(0) \delta A_i(t') \rangle_{\boldsymbol{\phi}(t)}$ represents the response of A_i at time t' due to a perturbation of ϕ^j at time t. Here, the angle brackets $\langle \cdots \rangle_{\boldsymbol{\phi}(t)}$ indicate an average over the equilibrium distribution for the instantaneous value of the field $\boldsymbol{\phi}(t)$. Thus the linear-response kernel $\chi_{ij}(t - t')$ can be represented by a matrix with time-dependent entries. Intuitively, the dynamic linear-response formula (2.59) captures the nonequilibrium response of an observable quantity A_j by integrating over its entire previous history of perturbations.

We can further simplify (2.59) by partially integrating to find

$$\langle \delta A_i(t) \rangle_{\text{neq}} \approx -\beta \int_{-\infty}^{t} \langle \delta f_j(0) \delta A_i(t - t') \rangle_{\boldsymbol{\phi}(t')} \dot{\phi}^j(t') \, \mathrm{d}t' \tag{2.60}$$

where $\dot{\phi}^j(t) \equiv \frac{\mathrm{d}}{\mathrm{d}t} \phi^j(t)$ is the velocity of ϕ^j at time t, and the boundary term vanishes if the system, for instance, started in an equilibrium state at $t \to -\infty$.

2.6.2 Generalized friction tensor

We now consider a system which is driven by a time-dependent control protocol Λ. The average power input to a system (the rate of work done on it) is the time derivative of (2.35), $\langle \mathcal{P} \rangle_{\Lambda(t)} = -\langle f_i(t) \rangle_{\text{neq}} \dot{\lambda}^i$, where the angle brackets $\langle \cdots \rangle_{\Lambda(t)}$ represent an average over the nonequilibrium ensemble at time t during control protocol Λ. Furthermore, the *excess*

power (rate of excess work) can be obtained by simply subtracting the equilibrium average,

$$\langle \mathcal{P}_{\text{ex}} \rangle_{\Lambda(t)} \equiv -\langle \delta f_i(t) \rangle_{\text{neq}} \dot{\lambda}^i \ . \tag{2.61}$$

Within the linear-response regime, the nonequilibrium conjugate-force fluctuations can be approximated using the dynamic linear-response formula (2.60):

$$\langle \delta f_i(t) \rangle_{\text{neq}} \approx -\beta \int_{-\infty}^{t} \langle \delta f_i(0) \delta f_j(t) \rangle_{\boldsymbol{\lambda}(t)} \dot{\lambda}^j \mathrm{d}t \ . \tag{2.62}$$

Expanding the control-parameter velocity in a Taylor series about its current value (at t) gives

$$\dot{\lambda}^j(t') \approx \dot{\lambda}^j(t) + \ddot{\lambda}^j(t'-t) + \mathcal{O}((t'-t)^2) \tag{2.63a}$$

$$\approx \dot{\lambda}^j(t) \ , \tag{2.63b}$$

where the approximation in (2.63b) is valid when $\dot{\lambda}^j \gg \ddot{\lambda}^j(t'-t)$ over time intervals $(t'-t)$ where the force response $\langle \delta f_i(0) \delta f_j(t) \rangle_{\boldsymbol{\lambda}(t)}$ is significantly greater than zero. Intuitively, this approximation requires that, over the timescales of conjugate-force relaxation, the control protocol has a velocity which is approximately constant.[15]

Substituting (2.63) in (2.62) allows us to take the control-parameter velocity $\dot{\lambda}^j$ out of the integrand. After then switching the integration variable $t - t' \to t''$ and flipping the integration bounds, we find that the average nonequilibrium conjugate force is

$$\langle \delta f_i(t) \rangle_{\text{neq}} \approx \zeta_{ij}(\boldsymbol{\lambda}) \dot{\lambda}^j(t) \ , \tag{2.64}$$

where ζ_{ij} is the (i,j)th element in the *generalized friction tensor*, given mathematically by the Green-Kubo relation [109]

$$\zeta_{ij}(\boldsymbol{\lambda}) = \beta \int_0^{\infty} \langle \delta f_i(0) \delta f_j(t) \rangle_{\boldsymbol{\lambda}} \, \mathrm{d}t \ . \tag{2.65}$$

The generalized friction is equivalent to the Kirkwood formulation of the friction tensor [114, 115]. Alternatively, the generalized friction can be decomposed into the product

$$\zeta_{ij}(\boldsymbol{\lambda}) \equiv \beta \langle \delta f_i \delta f_j \rangle_{\boldsymbol{\lambda}} \tau_{ij}^{\text{relax}}(\boldsymbol{\lambda}) \tag{2.66}$$

of the force covariance and the integral relaxation time $\tau_{ij}^{\text{relax}} \equiv \int_0^{\infty} \langle \delta f_i(0) \delta f_j(t) \rangle_{\boldsymbol{\lambda}} / \langle \delta f_i \delta f_j \rangle_{\boldsymbol{\lambda}} \mathrm{d}t$ [116]. Intuitively, this decomposition allows for a simple interpretation of the generalized friction in terms of the two factors, the variance of conjugate forces and the relaxation time.

[15] We will explore the limitations of this approximation in more depth in Chapter 6, with the particular motivation of how it can fail for stochastic control protocols.

Alternatively, Zulkowski and DeWeese [117] showed that—so long as the potential satisfies certain conditions on its asymptotic properties—the generalized friction is[16]

$$\zeta_{ij}(\boldsymbol{\lambda}) = \frac{1}{\beta D}\int \frac{\left[\partial_{\lambda^i}\Pi(x|\boldsymbol{\lambda})\right]\left[\partial_{\lambda^j}\Pi(x|\boldsymbol{\lambda})\right]}{\pi(x|\boldsymbol{\lambda})}\,\mathrm{d}x\ , \tag{2.67}$$

where $\Pi(x|\boldsymbol{\lambda}) \equiv \int_{-\infty}^{x}\pi(x|\boldsymbol{\lambda})\,\mathrm{d}x$ is the equilibrium cumulative distribution function, and D is the diffusion coefficient. This mathematical form of the generalized friction is preferable in many instances. For instance, if the potential energy function is known, then $\zeta_{ij}(\boldsymbol{\lambda})$ can be calculated from (2.67) by numerical integration, whereas if no analytical form of the force-autocovariance function is known (and very few are), then the friction from (2.65) must be determined by numerical simulation, which is much more computationally expensive.

Under the linear-response approximation, the average excess power at time t during control protocol Λ is given by the quadratic form

$$\langle \mathcal{P}_{\mathrm{ex}}\rangle_{\Lambda(t)} \approx \dot{\lambda}^i\zeta_{ij}(\boldsymbol{\lambda})\dot{\lambda}^j\ , \tag{2.68}$$

and the average excess work, integrated over the course of the protocol, is

$$\langle W_{\mathrm{ex}}\rangle_{\Lambda} = \int_0^\tau \langle \mathcal{P}_{\mathrm{ex}}\rangle_{\Lambda(t)}\mathrm{d}t \tag{2.69a}$$

$$\approx \int_0^\tau \dot{\lambda}^i\zeta_{ij}(\boldsymbol{\lambda})\dot{\lambda}^j\,\mathrm{d}t\ . \tag{2.69b}$$

Thus, the generalized friction allows us to estimate the nonequilibrium excess work using only equilibrium information about the system (the conjugate-force fluctuations).

2.6.3 Minimal-work control protocols

Historically, efforts to equip thermodynamic theories with geometric interpretations stretch back to Gibbs who sought to find "a general graphical method which can exhibit at once all the thermodynamic properties of a fluid concerned in reversible processes" [119]. Gibbs thought that the graphical representation of, for instance, constant-volume thermodynamic transformations as paths in a two-dimensional plane, provided the relatively abstract theory of classical thermodynamics with a more tractable interpretation. While Gibbs' analysis was restricted to reversible transformations, outside of thermodynamic equilibrium more recent efforts in the 1970s and 80s sought to define local metrics on the space of equilibrium thermodynamic states by using second derivatives of macroscopic quantities, the internal energy by Weinhold in [120] or the macroscopic entropy by Ruppeiner in [121], as well as first derivatives of the free energy with respect to intensive quantities by Schlögl in [122]. Such

[16] A similar mathematical form of a friction coefficient was also derived, in a different context, by Berezhkovskii and Szabo in [118].

interpretations were even used by Salamon *et al.* to understand minimal-dissipation control protocols in so-called step-equilibration processes [123] that bear striking resemblance with recent findings (see, for instance Ch. 7).

However, these formalisms can pose significant problems outside of the thermodynamic limit, where the metrics of Weinhold [120] and Ruppeiner [121] may not even exist. This train of thought was distilled by Crooks in 2007 by his introduction of the *thermodynamic length* $\mathcal{L}$ as an abstract measure of distance between equilibrium thermodynamic states [124]. Here, Crooks showed that the Fisher information $\mathcal{I}_{ij}(\boldsymbol{\lambda})$ can serve as a metric on the space of equilibrium states, and the minimum-dissipation control protocols are given by the geodesic curves connecting the initial and final equilibrium states.

The generalized friction provides a further refinement of this geometric view of thermodynamic space. By incorporating both the Fisher information matrix (through the force variance, as $\mathcal{I}_{ij}(\boldsymbol{\lambda}) \equiv \beta^2 \langle \delta f_i \delta f_j \rangle_{\boldsymbol{\lambda}}$) and the relaxation time (2.66), the generalized friction tensor equips the space of control parameters with a Riemannian manifold, defined by the metric $\zeta_{ij}(\boldsymbol{\lambda})$, and again having the interpretation of minimum-dissipation control protocols as geodesic curves.[17]

Using this interpretation, the excess work can be lower bounded with a Cauchy-Schwarz inequality

$$\langle W_{\mathrm{ex}} \rangle_{\Lambda} \geq \frac{\mathcal{L}^2}{\tau} \, , \tag{2.70}$$

where

$$\mathcal{L}(\Lambda) \equiv \int_0^{\tau} \sqrt{\dot{\lambda}^i \zeta_{ij}(\boldsymbol{\lambda}) \dot{\lambda}^j} \, \mathrm{d}t \tag{2.71}$$

is the *thermodynamic length* of protocol Λ, which is independent of the protocol duration τ.[18] This lower bound (2.70) is saturated for minimum-dissipation protocols, which are those that minimize the thermodynamic length of the control protocol. Furthermore, the minimum-dissipation protocol follows the geodesic on the space of control parameters. Additionally, by the Cauchy-Schwarz inequality, minimum-dissipation control protocols are those in which the integrand $\sqrt{\dot{\lambda}^i \zeta_{ij}(\boldsymbol{\lambda}) \dot{\lambda}^j}$ of the thermodynamic length is constant, and thus represent protocols which bear a constant excess power.

Throughout the thesis, we will often compare the predicted dissipation of a minimum-dissipation control protocol to that of a *naive protocol*, which is the particular control protocol that travels between the initial and final control parameter states at a constant

[17] As it turns out, there are mutually incompatible assumptions made with regard to the continuous limit when deriving Crooks' Fisher information metric or the generalized friction, but we will explore and elaborate on the details of these assumptions in Sec. 7.4.1.

[18] To recover Crooks' original measure of the thermodynamic length, we simply replace $\zeta_{ij}(\boldsymbol{\lambda})$ in $\mathcal{L}$ with the Fisher information matrix $\mathcal{I}_{ij}(\boldsymbol{\lambda})$.

velocity. Intuitively, the naive protocol serves as a first guess of the minimum-dissipation control protocol in the absence of any knowledge of the generalized friction $\zeta_{ij}(\boldsymbol{\lambda})$.

The identification of such paths in the general case requires solving the geodesic equation; however, for a single control parameter, minimum-dissipation (MD) protocols can be shown to satisfy the simple proportionality

$$\dot{\lambda}|_{\mathrm{MD}} \propto \frac{1}{\sqrt{\zeta(\lambda)}} \ , \tag{2.72}$$

where the constant of proportionality is fixed by enforcing that protocols be completed in a predetermined duration τ [109].

Intuitively, the one-dimensional result (2.72) states that, in order to minimize excess work, control protocols should move relatively slowly through regions of high generalized friction $\zeta(\boldsymbol{\lambda})$, and comparatively fast in regions of low friction. This formalism has been used to understand the minimal-work protocols in various model systems, including the inversion of magnetization in an Ising model [125], the optimal erasure of information [126], and several others [30, 65, 66, 76, 117, 125, 126, 127, 128].

In many of these papers, the minimum-dissipation protocols are referred to as *optimal protocols* because they are—in terms of the excess work—optimal. However, in the biomolecular context, it is not clear whether or not such protocols are in any sense optimal. While we suspect that energetically efficient molecular machines may have evolved through selective pressures based on limited energetic resources [129], it is exceedingly unlikely that the evolved molecular machines we observe *in vivo* today are the result of a single-variable optimization. More likely, they are an evolved result of trading off many distinct–and often conflicting–objectives [66, 130]. Thus, throughout this thesis, we will refer to protocols that minimize the excess work simply as minimum-dissipation (MD) protocols.

2.7 Model systems

Throughout this thesis, we make repeated use of a set of simple model systems to test our theoretical results. In this section, we introduce these models and discuss their relevance to true physical systems.

2.7.1 Harmonic trap

The simplest system that we use repeatedly is a Brownian particle diffusing in a harmonic potential

$$E(x|\lambda) = \frac{k_{\mathrm{trap}}}{2}(x-\lambda)^2 \ , \tag{2.73}$$

where k_{trap} is the trap strength, and the control parameter λ is the minimum of the potential. This simple potential represents, for instance, a colloidal particle diffusing in a optical trap [27]. In this case, the forces experienced by the diffusing particle are linear, $f =$

$-k_{\text{trap}}(x-\lambda)$, making further analytical investigation more simple. For example, the full time-dependent distributions of position $p(x,t)$ and work $p(W,t)$, among other things, have been solved exactly by Mazonka and Jarzynski for time-dependent control protocols [131].

As discussed in Sec. (2.6), the exact optimal protocols for the harmonic potential, traveling between λ_{i} at time $t = 0$ to λ_{f} at time $t = \tau$, can be found for both overdamped [106, 107] as well as underdamped particle dynamics [108]. For the exact results in overdamped systems, there are discrete jumps at the protocol boundaries [106, 107]. These jumps are suppressed in underdamped systems [108], as here such instantaneous changes in the control-parameter position require diverging input.

Using the linear-response approximation (Sec. 2.6.1), the generalized friction is simply equal to the viscous friction coefficient, $\zeta = \gamma$, and is therefore independent of the control-parameter value λ [109]. Thus, the minimum-dissipation protocol is predicted to have a constant velocity $\dot{\lambda}$. In contrast to the exact results, the optimal protocol predicted through the linear-response calculation contains no discontinuous jumps at the start and end of the protocol. However, in the exact solution [106], the magnitude of the jumps decreases as the protocol duration increases, and converges to the linear-response result as $\tau \to \infty$.

2.7.2 Periodic potential

Alternatively, we are often interested in the operation of cyclic molecular machines, and thus require a potential that satisfies periodic boundary conditions. A simple energy function that captures this is a cosine potential,

$$E(x,\lambda) = -\frac{E^{\ddagger}}{2}\cos 2\pi(x-\lambda) \ , \tag{2.74}$$

where $E^{\ddagger}$ is the height of the energy barrier separating a minimum from its periodic image, and λ is the location of the energy minimum. This potential can represent, for instance, a colloidal particle on a ring [27]. For $E^{\ddagger} \to \infty$, the generalized friction in the periodic potential (2.74) is the same as for (2.73), as the equilibrium distribution approaches a Gaussian; however, for finite $E^{\ddagger}$ the generalized friction—albeit still independent of λ—must be determined from (2.65).[19]

In both the model harmonic potential (2.73) and periodic potential (2.74), there is no underlying potential experienced by the system in the absence of control. While this is true, for instance, of a colloidal particle diffusing in solution, the same cannot be true for many biophysical systems, such as a molecular machine. To model the basic physics of molecular

[19]For periodic boundary conditions, the Zulkowski-DeWeese formula (2.67) cannot be used to calculate the generalized friction [128].

machines, we consider an underlying (λ-independent) periodic potential given by

$$E_{\text{mol}}(x) = -\frac{E^{\ddagger}}{2}\cos\frac{2\pi x}{\ell} \ , \tag{2.75}$$

which has a period of ℓ and a barrier of magnitude $E^{\ddagger}$ separating two adjacent minima. This potential can be considered on either a periodic x domain, or an infinite underlying x domain, representing either a rotary or translational molecular machine, respectively [132].

In either case, we impose a controlling potential $E_{\text{trap}}(x,\lambda)$ on top of the underlying potential, providing an additional influence on the preferred state of the system. A harmonic controlling potential does not obey any periodic boundary conditions, and thus can only be used in the absence of periodic boundary conditions. Here, the total potential experienced by the particle is

$$E_{\text{periodic}}(x,\lambda) = E_{\text{trap}}(x,\lambda) + E_{\text{mol}}(x) \tag{2.76a}$$

$$= \frac{k_{\text{trap}}}{2}(x-\lambda)^2 + \frac{E^{\ddagger}}{2}\cos\frac{2\pi x}{\ell} \ . \tag{2.76b}$$

In this case, the conjugate forces are still linear, $f(x) = k_{\text{trap}}(x-\lambda)$, but depending on the particular value of λ the dynamics of x can show distinct behaviours. For instance, if λ is located at a minimum of the periodic potential ($\lambda = 0, 1, 2, \dots$), then—for sufficiently large trap strength k_{trap}—the x-dynamics will look as if the system were subjected to a single harmonic potential, whereas if λ is at a peak in the underlying potential ($\lambda = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$) then the system will appear bistable, with two energetically equivalent minima, one on either side of the peak in $E_{\text{mol}}(x)$.

For this potential, no exact solution for the optimal protocol is known. However, the generalized friction can be determined using the Zulkowski-DeWeese formula (2.67) to find optimal protocols within the linear-response regime.

While we do not investigate the periodic underlying potential with periodic boundary conditions and the controlling potential in (2.74), [66] numerically investigated the trade-offs between work, accuracy, and speed in this system.

2.7.3 Fast-switching potential

Now, we consider a system which evolves on one of two potential energy functions $E_1(x,\lambda)$ or $E_2(x,\lambda)$. The particular potential experienced by the particle switches rapidly between the two different forms with rates that satisfy local detailed balance (2.20). In the limit of fast switching, the effective combined potential can be described by the *Potential of Mean Force* (PMF) of the two potentials, given mathematically by [133]

$$\beta E_{\text{PMF}}(x,\lambda) = -\ln\left[e^{-\beta E_1(x,\lambda)} + e^{-\beta E_2(x,\lambda)}\right] \ . \tag{2.77}$$

Such PMF-based models have been used, for instance, to explain the physics of the F_1 subcomponent of F_1F_o ATP synthase in electrorotation studies [134]. Here, the Kawaguchi-Sasa-Sagawa (KSS) model of F_1 constructs a PMF-based energy landscape in each of the three dominant angular orientations of the motor based on fast switching between two metastable conformational ensembles that interconvert on intermediate timescales. Specifically, the KSS potential is

$$\beta E_{\mathrm{KSS}}(\theta,\lambda) = \frac{k_{\mathrm{c}}}{2}(\theta-\lambda)^2 - \ln\left[e^{-k_{\mathrm{c}}\phi(\theta-\lambda)} + e^{\beta\Delta E_{\mathrm{S}}+\frac{\beta}{2}k_{\mathrm{c}}\phi^2} + \beta\Delta E_{\mathrm{S}}\right] \tag{2.78}$$

which is derived from (2.77) by substituting harmonic potentials for $E_1(x,\lambda)$ and $E_2(x,\lambda)$ (and discarding a physically immaterial overall energy offset). Here, the system coordinate θ is periodic, k_{c} is the strength of each of the individual harmonic potentials (assumed to be the same for both), ϕ is the angular offset between the two harmonic potentials, and ΔE_{S} is the energetic offset between their minima. While the KSS model violates the periodic boundary conditions—the underlying harmonic potentials do not respect periodic boundary conditions—it can be seen as a good empirical approximation for the parameters found from experiment. Specifically, the values for k_{c}, ϕ, and ΔE_{S} found in [134] are such that there is no significant probability flux across the periodic boundary. The KSS potential was used as a significant improvement over previous empirical reconstructions of the energy landscape of F_1 in each angular state, which required a 6th-order polynomial to capture the shape of the inferred distribution [135].

In Chapter 9 we use a similar model to explore the dissipation in a simple model of a mechanochemical machine; however, we use a PMF-based model that obeys the periodic boundary conditions, as the parameter ranges considered do not necessarily ensure that the probability remains localized near the dominant minimum in the potential. Specifically, we construct a PMF similar to (2.78) based on two sinusoidal potentials

$$E_1(\theta,\lambda) = \frac{E_1^{\ddagger}}{2}\cos 2\pi(\theta-\lambda) \tag{2.79a}$$

$$E_2(\theta,\lambda) = \frac{E_2^{\ddagger}}{2}\cos 2\pi(\theta-\lambda-\phi) + \Delta E_{\mathrm{S}}\ , \tag{2.79b}$$

where $E_1^{\ddagger}$ ($E_2^{\ddagger}$) is the amplitude of potential $E_1(\theta,\lambda)$ ($E_2(\theta,\lambda)$), while ϕ and ΔE_{s} are the respective angular offset and energetic shift of $E_2(\theta,\lambda)$ relative to $E_1(\theta,\lambda)$. Thus, the resulting PMF can be written as

$$\beta E_{\mathrm{PMF}}(\theta,\lambda) = -\ln\left[e^{-\frac{\beta}{2}E_1^{\ddagger}\cos 2\pi(\theta-\lambda)} + e^{-\frac{\beta}{2}E_2^{\ddagger}\cos 2\pi(\theta-\lambda-\phi)}\right]\ . \tag{2.80}$$

For certain values of the parameters $E^* \equiv E_2^\ddagger/E_1^\ddagger$, ϕ, and ΔE_S, the PMF in (2.80) is bistable, and thus represents a similar physical scenario as the KSS potential (2.78), while still satisfying the periodic boundary conditions of the angular coordinate θ.

Part I

Experimental tests of nonequilibrium theory

Chapter 3

DNA hairpins I: Calculating the generalized friction

3.1 Introduction

Reversible heat engines do not dissipate energy when operating infinitely slowly according to the Carnot cycle; their energetic efficiency is limited only by the entropy increase of the surroundings associated with the transfer of heat from a hot to a cold reservoir. In contrast, for engines operating irreversibly, the extra non-equilibrium energy cost associated with carrying out a process at a finite rate further reduces their efficiency [136]. This is the case of biological machines [35] that must operate under signaling, transport, and cell-cycle time constraints. For instance, F_oF_1-ATP synthase (section 1.1.2) can rotate up to ~ 350 revolutions/s [42]; the bacteriophage $\varphi 29$ packaging motor internalizes the 19.3 kbp viral genome into a small capsid at rates of 100 bp/sec, faster than the relaxation rate of the confined DNA [137]; and during sporulation, the *B. Subtilis* DNA translocase, SpoIIIE, transfers two thirds of its 4.2×10^6 bp genome between mother cell and pre-spore in only 15 minutes, i.e., at a transfer rate of nearly 4,000 bp/s [138]. The finite-time operations of these machines necessarily involve energy dissipation—often in the form of extra work—and it is of great interest to understand how they attain their large (over 70%) energetic efficiencies [53, 139].

DNA hairpins serve as a well-understood biomolecular model system in which to study the physics of strongly fluctuating nanoscale objects [140, 141]. Biochemically, hairpins are self-complementary sequences of single-stranded DNA or RNA and are a common secondary structure observed in biological systems. *In vivo*, hairpin structures serve a number of biological functions: they are known to bind to proteins that regulate transcription [142], appear as intermediates in recombination and replication [143], and present mechanical barriers to ribosomes during translation [144, 145]. Largely owing to their exceptional thermodynamic stability [146], hairpins also serve a number of *ex vivo* functions in nanotechnology, such as

fuel for autonomous DNA devices [147], or as a computational element in DNA-based logic circuits [148].

Technological advances in experimental methods have allowed hairpins to become a standard model system in single-molecule biophysics. In particular, high-resolution optical tweezers allow direct mechanical interaction with single hairpin structures, allowing for the observation and measurement of the biomolecular mechanics present in a nanoscale, strongly fluctuating system [149, 150]. To this end, hairpins have been used as a test-bed for theoretical predictions [24], and to further our understanding of mechanical forces, fluctuations, and dynamics at the nanoscale [141, 151, 152, 153, 154].

Broadly, hairpin structures can be classified by three parameters: the number of base-pairs in the stem of the hairpin structure (*stem length*), the number of base-pairs in the loop (*loop size*), and the percentage of stem base-pairs that are GC versus AT (or AU for RNA hairpins) (*stem GC content*) [141]. By tethering the base of the hairpin stem on either side—often through long strands of dsDNA—to two optical beads, optical tweezers can be used to directly apply forces to the hairpin stem, thus allowing a detailed study of hairpin physical properties. In the language of control theory (Section 2.6), the distance between the two optical trap foci is the control parameter λ.[1] Increasing λ, the optical tweezers apply mechanical forces to pull apart the hairpin into an elongated polymer, while decreasing λ compacts the polymer, favoring formation of the native hairpin structure. Between the all-folded and all-elongated conformational ensembles, the hairpin exhibits a regime of *hopping dynamics*, where the forces exerted by the optical trap to maintain a constant separation are bimodal, capturing the formation and destruction of the hairpin loop.

Here, we demonstrate experimentally the utility of the theoretical framework introduced in Sec. 2.6.2 for designing energetically efficient non-equilibrium processes, and propose that similar operation protocols may underlie the high efficiency observed in molecular machines. To this end, we subject single DNA hairpins to mechanical unfolding and refolding using protocols designed by this theory to minimize dissipation (Sec. 2.6.3), and we show that these protocols systematically and significantly reduce energy dissipation during the process. DNA hairpins are ideally suited for this test as the magnitude of the friction coefficient can be tuned by changing the molecule's length, the free energy difference, the free energy barrier, and the transition rates between its folded and unfolded states [141].

In Sec. 2.6.3 we showed that near equilibrium, the driving protocol $\lambda(t)_{\text{designed}}$ that minimizes the dissipation for a given total duration proceeds with a velocity proportional to the inverse square root of the local friction coefficient $\zeta(\lambda)$, $\mathrm{d}\lambda(t)_{\text{designed}}/\mathrm{d}t \propto \zeta(\lambda)^{-1/2}$ (2.72). The proportionality is fixed by the total duration of the protocol, so changing it corresponds

[1] In these experiments, we use the trap separation as the control parameter λ. However, it is also common, for instance, in [141], to use the applied force as the control parameter. The same theory can be applied in this case, but the calculation of the generalized friction and minimum-dissipation protocols is different.

to a global rescaling of all velocities. Other approaches to minimizing work [106, 107] require detailed knowledge of both the system's equilibrium energy landscape and non-equilibrium dynamics, and thus are experimentally challenging.

Throughout our subsequent discussion of hairpin pulling experiments, we refer to protocols that are predicted by the linear-response formalism discussed in Sec. 2.6.2 as *designed* protocols, as opposed to minimum-dissipation protocols. While the designed protocols are theoretically predicted to minimize the excess work in the linear-response regime, we have not shown that the protocols do in fact minimize excess work, only that they reduce excess work relative to a naive protocol that takes place over the same duration.

3.2 Experimental setup

We conducted high-resolution force separation measurements on a dual-trap optical tweezer setup using a solid-state 1064nm laser as outlined in [155].[2] The calibration procedure of the traps is detailed in [156]. The DNA tethers were formed between a 0.90-μm-diameter streptavidin-coated bead and a 1-μm-diameter antidigoxigenin-coated bead (Spherotech) held in separate optical traps. An oxygen scavenging system consisting of 100μg mL^{-1} glucose oxidase, 5 mg mL^{-1} dextrose (Sigma-Aldrich), 20μg mL^{-1} catalase (Calbiochem) was included in the buffer to prevent the formation of reactive singlet oxygen, thus increasing the lifetime of the DNA tethers.

To compare data from different experiments with the same hairpin sequence, we identify the trap extension $\lambda_{1/2}$ as the particular extension where the residence time in the folded and unfolded conformations is $\approx 50\%$. This extension provides a common reference extension to compare individual hairpins with the same sequence, that is independent of, for instance, variations in bead size. By reporting all subsequent extensions λ in a particular experiment relative to the $\lambda_{1/2}$ value determined for that particular molecule, we can easily compare data between experiments.

We selected hairpin DNA sequences that display hopping dynamics such that determining $\lambda_{1/2}$ was experimentally accessible. Very fast hopping dynamics are difficult to distinguish from noise, while very slow hopping dynamics require long data-acquisition times and laser exposure prior to the pulling experiments. Minimizing laser exposure avoids molecule photodamage. We investigated two hairpin sequences:

```
GAGTCCTGGATCCTGTTTTTTTTCAGGATCCAGGACTC          (3.1a)
TACCTGATCAGGTGCTTTTTTTTTGCACCTGATCAGGTA ,       (3.1b)
```

[2]The experiments were conducted primarily by Sara Tafoya (Scientific Application Development group, Lumicks, and formerly a PhD student at UC Berkeley), with help at the preliminary phase by Shixin Liu (The Rockefeller University, formerly a post-doctoral researcher at UC Berkeley).

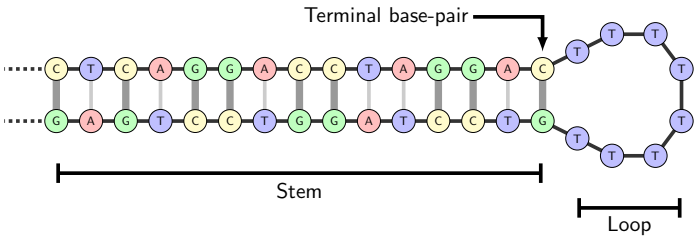

Figure 3.1: **Schematic of the fast-relaxing DNA hairpin sequence in its folded state.** General features of a DNA hairpin are the stem and loop. The stability of the stem is determined by both its length in base pairs (bp) and the percentage of the base pairs within the stem that are GC. Both the fast- and slow-relaxing hairpin sequences (3.1) contain a 15-bp stem sequence and an 8-base loop sequence. The terminal base pair is the last base pair in the native hairpin structure stem before the loop sequence.

the first sequence (3.1a) has a relaxation time of $t_{1/2} \approx 240$ milliseconds at $\lambda_{1/2}$ [141], while the second sequence (3.1b) has a faster relaxation time of $t_{1/2} \approx 30$ milliseconds at $\lambda_{1/2}$. Due to this difference in the relaxation times at $\lambda_{1/2}$, in the remainder of the thesis we will refer to the respective sequences as the fast-relaxing hairpin (3.1b) and the slow-relaxing hairpin (3.1a).

Biochemically, the difference between the fast- and slow-relaxing hairpins is a relative increase in the fraction of GC base pairs in the fast-relaxing hairpin stem at the loop neck. This change in sequence is expected to facilitate the nucleation of the native hairpin structure and avoid molecule misfolding [157]. The stability of a particular DNA hairpin sequence depends on the presence of GC vs. AT base pairs in the stem due to the different strength of these base pair configurations [141]. Figure 3.1 shows a schematic depiction of the fast-relaxing hairpin in its folded state, with the different-width vertical bonds indicating the relative strengths of GC and AT base pairs.

While both of these sequences exhibit hopping dynamics appropriate for our experiments, the bulk of our analysis is focused on the fast-relaxing sequence (3.1b) as this particular sequence allowed experimental analysis of a wider range of protocol durations. For instance, in the slow-relaxing hairpin sequence, for protocol durations $\tau < 1/2$ seconds, a significant number of refolding protocols did not show refolding of the hairpin structure over the course of the protocol. We will revisit the results from the slow-relaxing hairpin in Chapter 5.

Variation in the bead sizes, small differences in chemical attachments, and non-specific interactions of the hairpin sequence with the bead surface can lead to molecule-to-molecule variation. We minimized the contribution of trap distance variation by subtracting the value of $\lambda_{1/2}$ from all trap separation measurements. However, other unaccounted sources, such as errors in the stiffness calibration and natural variation in the molecules' persistence length also contributed to molecule-to-molecule variation in the unfolding/refolding trajectories. For instance, most commercially available beads have a root-mean-squared (RMS) variation

in radius of $3-6\%$, which (assuming a 4% error in reported bead sizes and using individual calibration measurements for each bead pair) would produce an error in stiffness calibration of $\approx 4\%$. Additionally, the standard deviation in persistence-length measurements of single-stranded DNA can be as high as 17%, indicating that experimental variation can have a large impact on the measured physical properties of DNA [158].

3.3 Equilibrium sampling

Each molecule was initially probed to find $\lambda_{1/2}$: the distance between the optical trap centers was gradually increased until the residence time at the folded and unfolded conformations is $\approx 50\%$. On identification of $\lambda_{1/2}$, a systematic error of ≈ 2.5 nm was introduced as a small difference of a few millivolts (mV) between the instructions sent by the computer and the analog signal received by the steering mirror of the trap. This problem is not present when measuring changes in separation, because, in calculating relative distances, the offset is canceled. However, we estimated the increase of the theoretically predicted excess work of the offset protocols relative to their centered values, and found that such an error should only lead to a cycle work overestimate of $\approx 6\%$.

After determining $\lambda_{1/2}$, we sampled 15 different fixed relative trap separations $\lambda - \lambda_{1/2}$ between a minimum at $\lambda - \lambda_{1/2} = -50$ nm and a maximum at $\lambda - \lambda_{1/2} = 50$ nm. We measured the equilibrium force fluctuations in 10-nm increments far from $\lambda_{1/2}$ and in 5-nm increments near $\lambda_{1/2}$ so as to better resolve the friction variation in the hopping regime. To extract the relevant information from the raw experimental data, the time series are segmented into individual fixed-extension sections, as shown in Fig. 3.2a. Each individual molecule provides a sample of equilibrium force fluctuations at each of the 15 trap separations.[3]

Figure 3.2a shows that, for very small or very large trap separations, the force fluctuates around a single mean value corresponding to the folded or unfolded configuration. Conversely, for intermediate trap separations, the force fluctuates between two different values, reflecting the hopping dynamics of the DNA hairpin, sampling the folded and unfolded configurations. For each separation, we calculated the probability distribution of equilibrium forces (Fig 3.2b), which are unimodal for small and large trap separations, and bimodal for intermediate separations. For each trap separation λ we also calculated the force autocovariance function $\langle \delta f(0) \delta f(t) \rangle_\lambda$ (Fig. 3.2c) which, as expected, in the hopping regime has a force variance—the $t = 0$ value of the force autocovariance—that is larger and has fluctuations which decay more slowly than when an extreme trap separation holds the hairpin in a single conformation.

[3] A single molecule might not give data for all 15 extension values if, for instance, the hairpin broke midway through data acquisition. Such hairpin breakage events, however, can be easily identified by a large instantaneous drop in the forces being applied by the optical traps, and thus do not corrupt the resulting analysis.

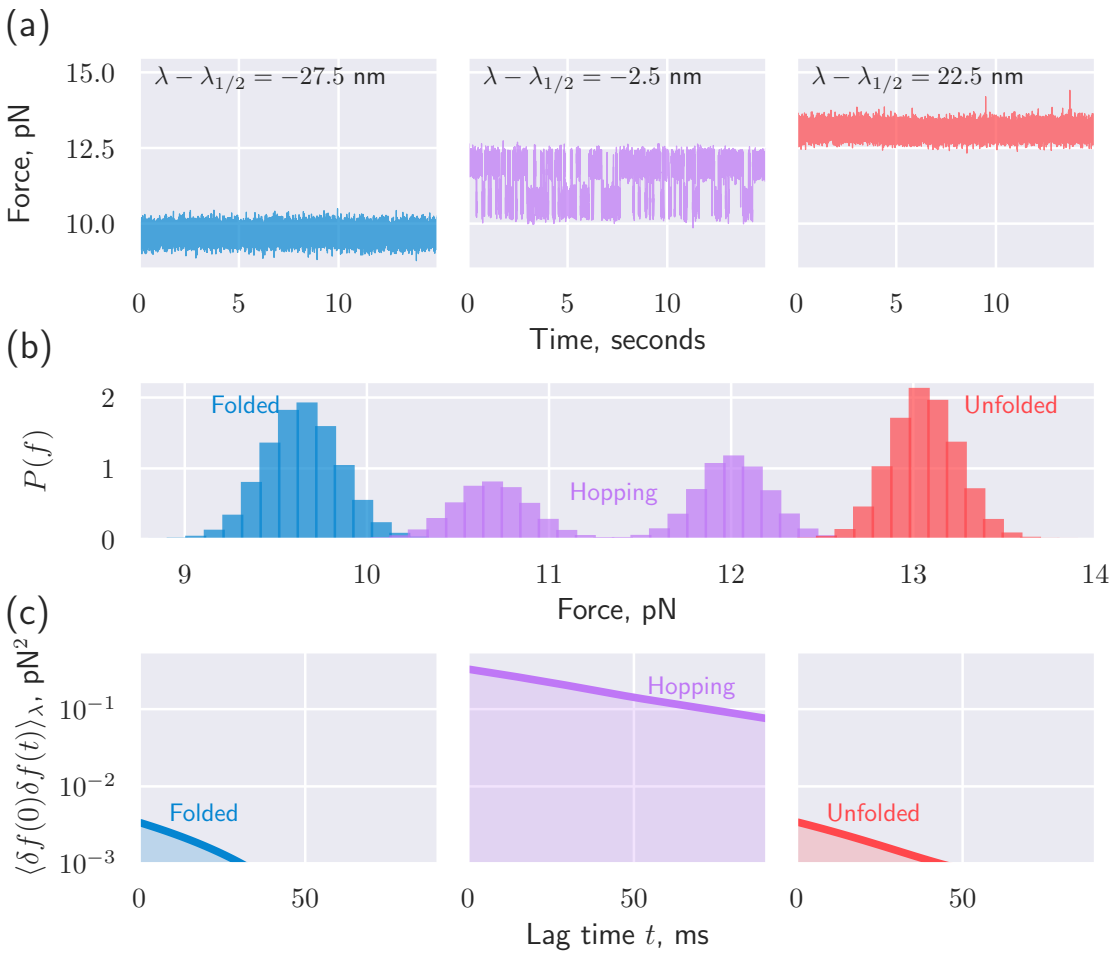

Figure 3.2: **Equilibrium sampling of DNA hairpin.** (a) Sample traces as a function of time for folded hairpin (left, blue), hopping hairpin (center, purple), and unfolded hairpin (right, red). (b) Equilibrium force distributions and (c) equilibrium force autocovariances as a function of lag time for corresponding fixed optical trap separations.

3.3.1 Estimating the generalized friction coefficient

We used the equilibrium data obtained from 20 different molecules to infer the generalized friction (2.65) of the rapidly-relaxing hairpin. For each fixed trap separation, we had ≈ 20 time-series of equilibrium fluctuation data from which to estimate the force-autocovariance function. Each data series gives a single autocovariance function and thus permits a single estimate of the generalized friction, so the set of 20 molecules gives a distribution of friction estimates.

In order to infer the mean friction and standard error of the mean (SEM) at each relative extension, we use jack-knife resampling [83], where we define a set of resampled mean estimates:

$$\overline{\zeta}^{(i)}(\lambda) = \frac{1}{N-1} \sum_{j=1, j \neq i}^{N} \zeta_j(\lambda) \ . \tag{3.2}$$

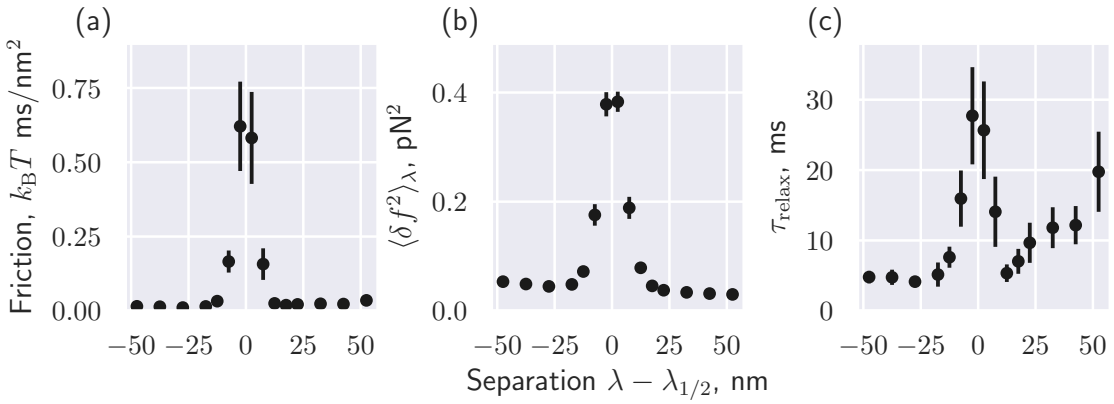

Figure 3.3: **Generalized friction as a function of trap separation in a DNA hairpin**. (a) Generalized friction coefficient $\zeta(\lambda)$, (b) force variance $\langle \delta f^2 \rangle_\lambda$, and (c) integral relaxation time as a function of fixed optical trap separation.

obtained by omitting the ith observation from the distribution of $\zeta(\lambda)$ values. N is the total number of samples for a given extension. The estimate of the mean friction is then obtained by averaging over the resampled estimates as

$$\overline{\zeta}(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \overline{\zeta}^{(i)}(\lambda) \ . \tag{3.3}$$

The jack-knife mean estimate (3.3) gives the same result as the typical sample mean $\overline{\zeta}(\lambda) = \frac{1}{N} \sum_j \zeta_j(\lambda)$, but also allows us to estimate the variance of our estimate of the mean via the distribution of resampled means $\overline{\zeta}(\lambda)_i$ (3.2):

$$\left\langle \delta \overline{\zeta}^2 \right\rangle = \frac{N-1}{N} \sum_{i=1}^{N} \left[\overline{\zeta}^{(i)}(\lambda) - \overline{\zeta}(\lambda) \right]^2 \ . \tag{3.4}$$

Here, $\delta\overline{\zeta} \equiv \overline{\zeta}^{(i)} - \overline{\zeta}$ is the difference between the ith resampled estimate of the generalized friction and the estimate of the mean (3.3). The standard error of the mean is then simply the square-root of the estimator variance (3.4) [83].[4]

Figure 3.3 shows the estimates of the generalized friction (Fig. 3.3a), force variance (Fig. 3.3b), and integral relaxation time (Fig. 3.3c) as a function of optical trap separation λ. The force variance peaks at $\lambda_{1/2}$, where the hairpin spends roughly equal time in the open and closed configurations. Likewise, the integral relaxation time peaks at $\lambda_{1/2}$, reflecting that, to equilibrate, the hairpin must relax across the barrier separating the folded and

[4]While for large sample sizes, resampling techniques like the jack-knife method should recover the same estimates as more standard calculations of the standard error of the mean, resampling techniques can behave differently for small numbers of samples (such as we had for our estimates of the generalized friction coefficient) and often perform better than the standard normal-distribution approximation [159].

unfolded states. At room temperature, a 1-μm bead experiencing Stokes drag (with friction $\gamma = 6\pi\eta R$ for water viscosity η and bead radius R) in water and confined by a $k = 0.25$-pN/nm optical trap has a relaxation time $\gamma/k \sim 180$ μs, an order of magnitude below the minimum observed relaxation time for the entire hairpin construct, indicating that the beads do not significantly impact the relaxation times (typically in the low milliseconds) observed in experiments. Finally, the generalized friction coefficient (2.66)–the product of force variance and integral relaxation time–also peaks at $\lambda_{1/2}$.

3.3.2 Designing reduced-dissipation control protocols

As mentioned in Section 3.1—and elaborated upon in 2.6.2—linear-response theory predicts that (near equilibrium) the minimum-dissipation protocol proceeds with a pulling speed–or velocity of the steering trap–that scales as the inverse square root of the friction coefficient $\zeta(\lambda)$ [109]: pulling fast at extreme separations, where the friction coefficient is small, and slow around $\lambda_{1/2}$, where friction peaks. Intuitively, a slow velocity near $\lambda_{1/2}$ provides more time for thermal fluctuations to induce the unfolding or folding of the DNA hairpin without additional work input, thereby decreasing the work required to drive the DNA hairpin between conformations [127].

Predicting the minimum-dissipation protocol from the experimental data requires simply calculating $1/\sqrt{\zeta(\lambda)}$ and rescaling the velocities globally so as to enforce the constraint of a fixed protocol duration (Fig. 3.4a). To ease the experimental implementation, we fit the empirically determined optimal velocities to a piecewise-constant acceleration profile.

Qualitatively, the empirically estimated designed-protocol velocity in Fig. 3.4a is constant near $\lambda_{1/2}$ and far away from $\lambda_{1/2}$, so we apply these constraints to all considered models, interpolating between these regions with constant accelerations. Here, the model is parameterized by the constant-velocity region boundaries and the constant velocities themselves. In particular, the general model is defined as

$$\left.\frac{\mathrm{d}\lambda}{\mathrm{d}t}\right|_{\text{designed}} = \begin{cases} \gamma_{\mathrm{L}} & \text{for} \quad \lambda \in [-50nm, \alpha_{\mathrm{L}}] \\ \omega_{\mathrm{L}}(\lambda - \lambda(\alpha_{\mathrm{L}})) + \gamma_{\mathrm{L}} & \text{for} \quad \lambda \in [\alpha_{\mathrm{L}}, \beta_{\mathrm{L}}] \\ \gamma_{\mathrm{C}} & \text{for} \quad \lambda \in [\beta_{\mathrm{L}}, \beta_{\mathrm{R}}] \\ \omega_{\mathrm{R}}(\lambda - \lambda(\beta_{\mathrm{R}})) + \gamma_{\mathrm{C}} & \text{for} \quad \lambda \in [\beta_{\mathrm{R}}, \alpha_{\mathrm{R}}] \\ \gamma_{\mathrm{R}} & \text{for} \quad \lambda \in [\alpha_{\mathrm{R}}, 50nm] \end{cases} \tag{3.5}$$

where $\gamma_{\mathrm{L/C/R}}$ are protocol velocities for the constant-velocity sections in the left/center/right region of the plot, α_{L} (α_{R}) is the point at which the left (right) constant-velocity section ends (begins), $\beta_{\mathrm{L/R}}$ are the positions at which the central constant-velocity section (near $\lambda_{1/2}$) starts/ends, and $\omega_{\mathrm{L/R}}$ are the slopes of the left/right constant-acceleration segments. The constant-acceleration slopes $\omega_{\mathrm{L/R}}$ are not independent parameters, but are determined

by the other parameters as

$$\omega_{\mathrm{L}} = \frac{\gamma_{\mathrm{C}} - \gamma_{\mathrm{L}}}{\beta_{\mathrm{L}} - \alpha_{\mathrm{L}}} \tag{3.6a}$$

$$\omega_{\mathrm{R}} = \frac{\gamma_{\mathrm{R}} - \gamma_{\mathrm{C}}}{\alpha_{\mathrm{R}} - \alpha_{\mathrm{C}}} \ . \tag{3.6b}$$

In the most general case, this model has 7 parameters: $\{\gamma_{\mathrm{L}}, \gamma_{\mathrm{C}}, \gamma_{\mathrm{R}}, \alpha_{\mathrm{L}}, \alpha_{\mathrm{R}}, \beta_{\mathrm{L}}, \beta_{\mathrm{R}}\}$. Different models can be obtained from the velocity profiles by imposing different symmetries, such as inversion symmetry about $\lambda_{1/2}$, thus reducing the number of free parameters.

Model selection and information criterion

To judge the relative quality of different models with varying numbers of parameters, we compare them based on their Akaike Information Criterion (AIC). The AIC is an information measure used in model selection to discriminate between different models with varying numbers of free parameters, so as to avoid overfitting [160]. Mathematically, the AIC of a given model $M(\theta_n)$ with N parameters θ_n given a set of data $X = \{x_i\}$ is [161]

$$\mathrm{AIC}(X|M) = -\ln \mathcal{L}_{\max}(X|M) + 2N \ , \tag{3.7}$$

where $\mathcal{L}_{\max}(X|M)$ is the maximum likelihood of the data X given the model M, maximized over the parameters θ_n [83]. The $2N$ term penalizes models with a large number of parameters, providing a trade-off between a good fit to the data and overfitting to noise.[5] The model which minimizes the AIC is the preferable parameterization of the data. The particular model selected by minimization of the AIC is symmetric in the α and β parameters, but asymmetric in the magnitudes of the constant-velocity sections ($\gamma_{\mathrm{L}} \neq \gamma_{\mathrm{R}} \neq \gamma_{\mathrm{C}}$).

For a protocol duration of $\tau = 1/8$ seconds—the fastest protocol duration used in our following analysis of the fast-relaxing hairpin—Figure 3.4 shows the optimal-velocity profile (Fig. 3.4a) and the corresponding optimal protocols as a function of the scaled protocol time t/τ (Fig. 3.4b), which differ substantially from naive protocols (proceeding with a constant velocity) that are completed in the same duration. In particular, instantaneous driving velocities varied by a factor of approximately six within a given designed protocol.

This analysis shows that DNA hairpins, despite their simplicity as a model system for single-molecule biophysics, show a diverse set of behaviors as a function of extension. Furthermore, mapping this variation onto the generalized friction coefficient predicts minimum-dissipation unfolding/refolding protocols which are markedly different than naive protocols.

[5] A related measure, the *Bayesian Information Criterion* (BIC), is given by $\mathrm{BIC}(M) = -\ln \mathcal{L}_{\max}(M|X) + N \ln n_x$ where n_x is the number of observations. There are subtle differences between the two measures, with the BIC more heavily penalizing complex models, and thus generally preferring simpler models than the AIC [160]. However, for the cases we considered, AIC and BIC preferred the same model.

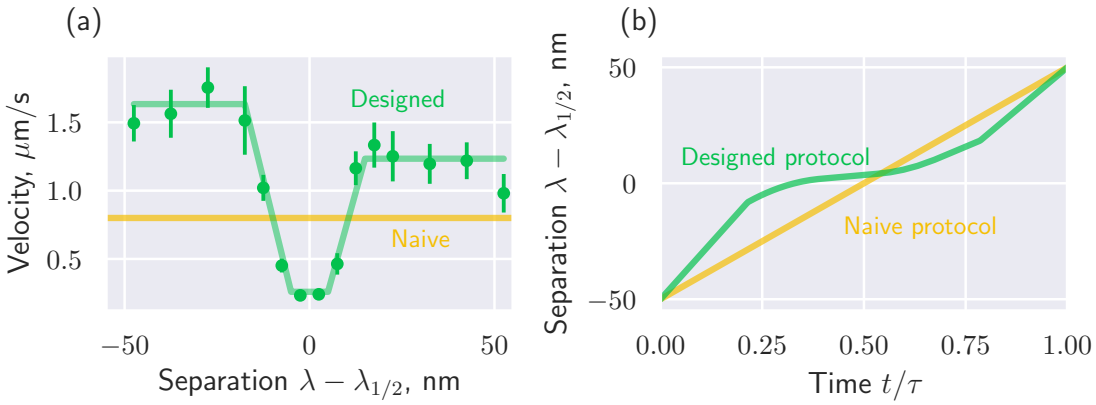

Figure 3.4: **Designed protocols in a DNA hairpin**. (a) For a protocol duration of 1/8 seconds, the designed velocity $\mathrm{d}\lambda/\mathrm{d}t \propto \zeta^{1/2}(\lambda)$ (green points) with best-fit model (green curve) that minimizes the Akaike information criterion [161], compared with naive velocity (yellow curve). (b) Designed (green) and naive (yellow) protocols as functions of t/τ. (Designed and naive velocities scale inversely with protocol duration τ.)

In Ch. 4, we implement both naive and designed unfolding/refolding control protocols in single DNA hairpins and directly measure the energetic costs of implementing each control strategy.

Chapter 4

DNA Hairpins II: reducing dissipation in nonequilibrium protocols

To test the utility of the generalized friction to reduce the dissipation in nonequilibrium processes, we implemented the designed protocols predicted in Chapter 3 for various protocol durations, comparing the dissipation in designed protocols to naive, constant-velocity protocols. Each hairpin control protocol gives a single force-separation curve to analyze. Thus by performing each type of pulling protocol (folding or unfolding, designed or naive) many times, we build up a statistical sample to perform analysis on. Unfolding protocols start (at $t = 0$) at low separations and end (at $t = \tau$) at high separations, while refolding protocols start at high separations and end at low separations.

The qualitative shape of individual force-separation curves (at equilibrium) is well-known from previous single-molecule studies [152, 162]. Broadly, the force-separation curve contains three regions: the low-separation regime ($\lambda < \lambda_{1/2}$), the hopping regime ($\lambda \approx \lambda_{1/2}$), and the high-separation regime ($\lambda > \lambda_{1/2}$). In the low- and high-separation regimes, the force monotonically increases with separation, while in the hopping regime, the force-separation curve shows signatures of 'force-rip' events [163] where the hairpin (see Fig. 3.1) opens or closes. Physically, when the hairpin opens, the length of ssDNA that is stored in the hairpin loop is liberated, leading to a longer effective contour length of polymer between the optical beads. Thus, opening of the hairpin manifests on the force-separation curve as a sharp decrease in the force at a particular separation. Conversely, the re-formation of the hairpin is indicated on the force-separation curve by a sharp increase in the force at a particular separation.

As the protocol duration is decreased—and therefore mean protocol velocity is increased—the number of force-rip events decreases, and the location of the force rips shifts to higher forces for unfolding protocols and lower forces for refolding protocols. For quasistatic protocols—arising in the $\tau \to \infty$ limit—the force-extension curves of unfolding and refolding protocols

are statistically indistinguishable. As the protocol duration shortens, the hairpins are pushed out of equilibrium during the protocol, and the unfolding and refolding protocols become statistically distinct populations.

Figure 4.1 shows sample force-extension curves for unfolding and refolding protocols in a DNA hairpin. For long protocol durations, in both the unfolding and refolding protocols, there are several force-rip events near a separation of $\lambda - \lambda_{1/2} = 0$, while for decreasing protocol duration—and therefore increasing mean protocol speed—the number of force-rip events decreases, and the location of the force-rip event shifts towards the end of the protocol, indicating that in each case the hairpin is significantly out of equilibrium and lagging behind its equilibrium state.

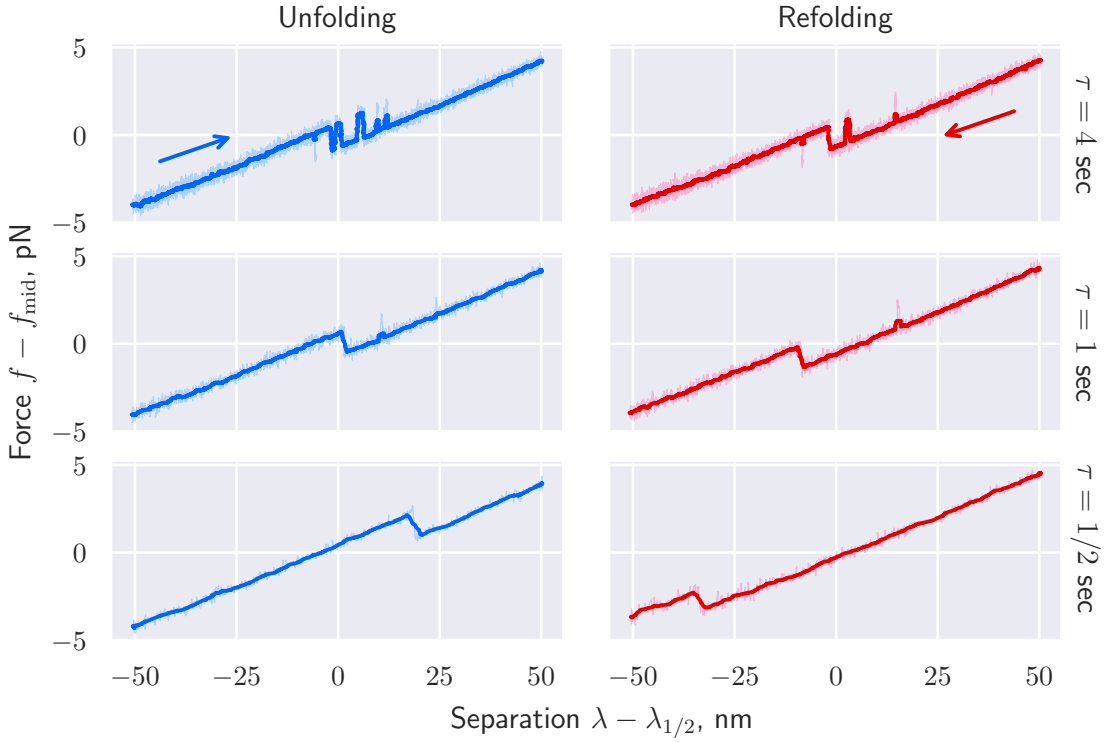

Figure 4.1: **Sample force-separation curves for DNA hairpins at different pulling speeds**. Force-separation curves for (blue, left) unfolding and (red, right) refolding protocols, at three different protocol durations $\tau = 4$, 1, or 1/2 sec (rows). Light curves show the raw experimental data, and dark curves show the same data smoothed with a Savitsky-Golay filter. Arrows indicate the temporal direction of the protocols.

The following sections outline the analysis performed on the fast-relaxing hairpin construct, for four protocol classes: designed unfolding, designed refolding, naive unfolding, and naive refolding. We performed each protocol for six different durations, $\tau = 1/8$, 1/4, 1/2,

1,2, and 4 seconds,[1] with each duration including experiments performed on 14, 9, 8, 8, and 10 separate hairpin molecules, and resulting in 888, 590, 396, 590, 592, and 472 individual protocol realizations (for each protocol class), respectively. From these data sets, we extract a series of physical quantities to characterize and quantify the physical properties and dissipation in the system.

4.1 Unfolding/refolding force identification

The first quantity that we discuss is the unfolding (refolding) force of a particular trajectory. The force at which the hairpin unfolds can provide a useful measure of the mechanical properties of the particular hairpin sequence. For instance, the unfolding force can be used to infer kinetic rate constants [163], as well as the location of the transition state and the activation free energy [164]. For our purposes, we make use of the unfolding/refolding force as a proxy for how far out of equilibrium the hairpin is along a particular protocol. In particular, for an unfolding (refolding) protocol that takes a finite duration, the unfolding (refolding) force will be increased (decreased) relative to its quasistatic value, regardless of how the unfolding force is defined. Ultimately this arises from the inability of the hairpin to completely sample its configuration space at each time in the time-dependent protocol, and thus the force lags behind its equilibrium value at the corresponding trap separation.

Figure 4.2 shows a pair of designed and naive protocols with the unfolding and refolding forces identified. Intuitively, the unfolding (refolding) force is, simply, the force at which the hairpin unfolds (refolds) along a particular protocol. More specifically, for our purposes, the unfolding (refolding) force is the maximum (minimum) force exerted by the optical tweezers before a force-rip event. There is, however, a subtlety in how exactly to define the unfolding/refolding force. For force-separation curves like those depicted in the $\tau = 1/2$ sec row of Fig. 4.1, the unfolding/refolding force is unambiguous; however, for force-extension curves in which the hairpin loop unfolds and refolds multiple time along a single protocol—such as in the $\tau = 4$ sec row of Fig. 4.1—the definition is less clear.

For a particular realization of an unfolding or refolding protocol, the force-separation curve $f(\lambda)$ captures the stochastic system response. In general, for any particular unfolding (refolding) force-separation curve, each force-rip event provides a candidate unfolding (refolding) force f^{U} (f^{R}), the local maximum (minimum) of the force-separation curve in its immediate vicinity. Thus, for each force-separation curve, we obtain a set of candidate unfolding (refolding) forces $\mathbb{F}_{\mathrm{U}} = \{f_i\}$ ($\mathbb{F}_{\mathrm{R}} = \{\tilde{f}_i\}$) where each $f(\lambda_i)$ ($\tilde{f}_i$) is a single candidate unfolding (refolding) force.

[1] These were the target durations of each protocol; however, noise in the instrument led to small variations in the actual protocol durations completed. In particular, the mean protocol durations were $\tau = 0.13, 0.24, 0.48, 0.93, 1.8$, and 3.7 seconds. Figures showing trends as a function of protocol duration will use the actual mean protocol durations, but for simplicity the text will refer to the target durations.

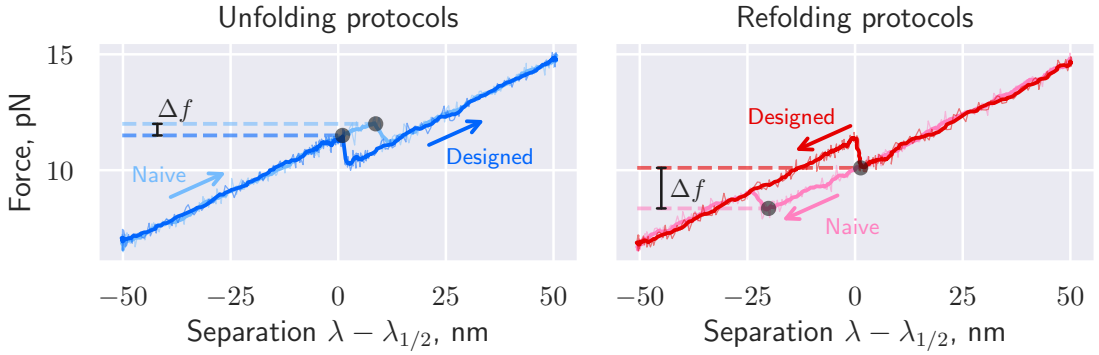

Figure 4.2: **Unfolding/refolding force identification**. (A) Example force-separation curves from a sample molecule for protocol duration $\tau = 0.13$ s, highlighting the unfolding (left) and refolding (right) events (black dots) and the corresponding force (dashed lines) for designed (dark blue and red) and naive (light blue and pink) protocols.

We define the unfolding (refolding) force f^{U} (f^{R}) as the maximum (minimum) of $\mathbb{F}_{\mathrm{U}}$ ($\mathbb{F}_{\mathrm{R}}$):

$$f^{\mathrm{U}} \equiv \max_{\mathbb{F}_{\mathrm{U}}} [f_i(\lambda)] \tag{4.1a}$$

$$f^{\mathrm{R}} \equiv \min_{\mathbb{F}_{\mathrm{R}}} [f_i(\lambda)] \ . \tag{4.1b}$$

Our definition of the unfolding (refolding) force f^{U} ($\tilde{f}^{\mathrm{R}}$) ensures that it lies on the boundary of one of the force-rip regions of the force-extension curve. Other definitions of the unfolding/refolding forces include swapping the maximum and minimum in (4.1), which identifies the unfolding force as the maximum force before the *first* unfolding event, and the refolding force as the minimum force before the first refolding event [163].

Intuitively, minimizing the average work often amounts to keeping the driven system as close to equilibrium as possible throughout the transformation. Thus—given that the unfolding (refolding) forces serve as a proxy for how far out of equilibrium the system is—we expect that designed protocols (Ch. 3) should have, on average, lower unfolding forces and higher refolding forces than their naive counterparts. We analyze the trend of average unfolding force differences between designed and naive protocols over different average speeds (inverse of protocol durations). Over the range of durations from $\tau = 1/8 \to 4$ sec, the force-separation curves display significant differences in the unfolding and refolding forces. Figure 4.3a shows the distributions of unfolding force differences $f^{\mathrm{U}}_{\mathrm{naive}} - f^{\mathrm{U}}_{\mathrm{designed}}$ and refolding force differences $f^{\mathrm{R}}_{\mathrm{naive}} - f^{\mathrm{R}}_{\mathrm{designed}}$ for three different protocol durations. As predicted by theory, on average, the DNA hairpin unfolded at lower forces and refolded at higher forces during designed protocols than during naive protocols of the same duration. This is indicated by the skewing of the force-difference distributions away from zero, towards

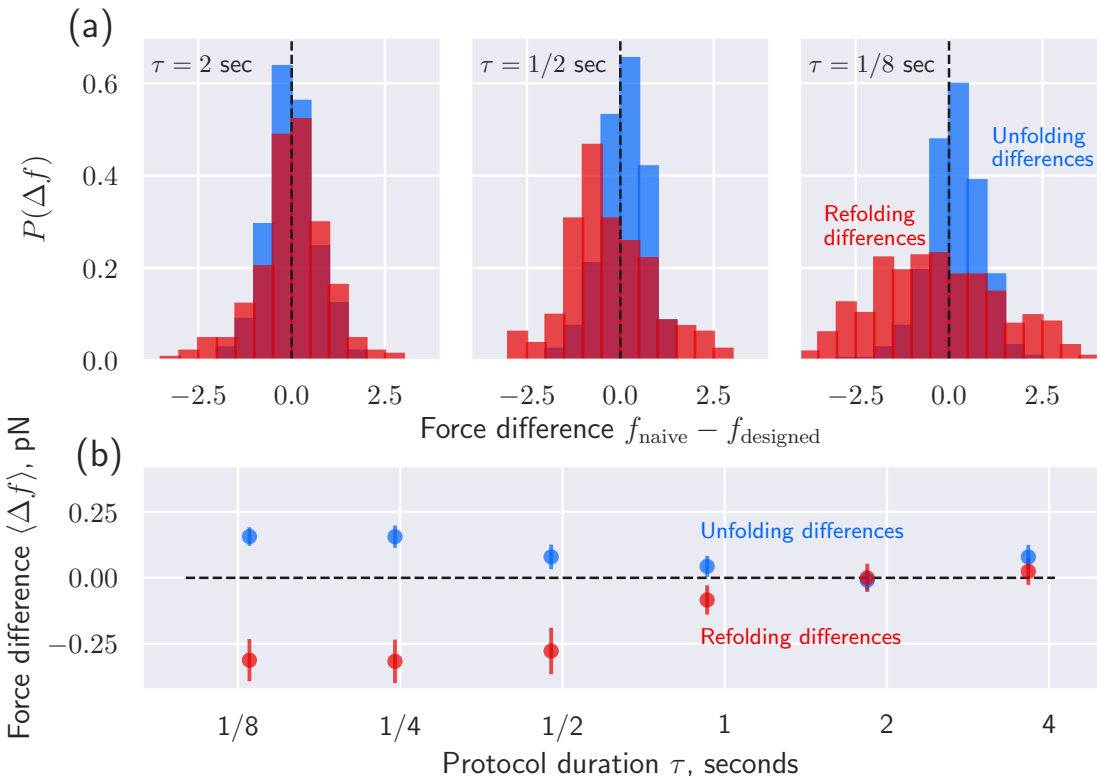

Figure 4.3: **Unfolding force distributions for designed and naive nonequilibrium protocols**. (a) Distributions of differences $f_{\text{naive}} - f_{\text{designed}}$ between naive and designed unfolding (blue) and refolding (red) forces. (b) Mean and standard error for unfolding and refolding force differences as a function of protocol duration. On average, the designed protocol unfolds at lower force and refolds at a higher force than the corresponding naive protocol.

positive (negative) values for unfolding (refolding) differences. Furthermore, the magnitude of the mean force difference is greater for faster protocols (Fig. 4.3b).

These results imply that the designed protocols display a lower hysteresis than naive protocols, a trend that is more prominent in faster protocols where the system is driven farther from equilibrium.

In Figure 4.3 we perform our analysis on the unfolding (refolding) force differences $\Delta f = f^{\mathrm{U}} - f^{\mathrm{R}}$—as opposed to the raw unfolding (refolding) force values f^{U} (f^{R})—for practical reasons. In comparing raw unfolding force between different molecules, the resulting analysis is affected by inter-molecule variation due to the instrument calibration. In particular, factors such as bead-size variation and nonspecific interactions of the dsDNA handles with the beads manifest as errors in the force offset (or, in comparing between molecules, the average force at $\lambda = -50$) for a particular molecule. Thus, the raw folding-force distributions would have erroneously large variance, and it would be much more difficult to identify statistically meaningful differences between the different protocol classes. However,

by comparing pairs of protocols from the same molecule (and thus the same force offset), the variation due to such experimental factors is canceled, resulting in a much more precise quantification of the protocol-specific differences in unfolding/refolding forces. Similarly, one could look at alternative measures that achieve the same cancellation, such as the *force hysteresis*, obtained for each protocol class by subtracting the unfolding force and refolding force of the same protocol type within a particular molecule: $f^{\text{hyst}}_{\text{naive}} \equiv f^{\text{U}}_{\text{naive}} - f^{\text{R}}_{\text{naive}}$ and $f^{\text{hyst}}_{\text{designed}} \equiv f^{\text{U}}_{\text{designed}} - f^{\text{R}}_{\text{designed}}$. Appendix B.1 further elaborates on these alternative force measures and their results.

4.2 Excess work measurements

Next, we focus on calculating the average excess work required to perform designed or naive protocols, thus closing in on the central question of these experiments: what is the utility of the generalized friction in designing protocols that reduce dissipation (excess work) out of equilibrium? For a given force-extension curve, the work required during a particular realization of the protocol Λ is the integral

$$W = \int_{\lambda(0)}^{\lambda(t)} f(\lambda)\mathrm{d}\lambda \ . \tag{4.2}$$

For a given force-extension curve, with forces f_i and separations λ_i, at time point t_i with $i \in [1, N]$, we approximate the 'true' work in (4.2) by numerically integrating the force-extension curve with a trapezoidal rule

$$W = \sum_{n=2}^{N} \frac{f_i + f_{i-1}}{2}(\lambda_i - \lambda_{i-1}) \ . \tag{4.3}$$

Each individual realization provides a numerical value for the protocol work, drawn randomly from the work distribution of the particular protocol class. However, naively comparing the raw work values for designed and naive protocols across different molecules can produce results that contain several experimental artifacts. In particular, errors in the instrument calibration manifest as a shift in the force offset. Even a small error in this force, when integrated over an entire protocol, can lead to significant differences in the work distributions of different molecules. For instance, a molecule with a force offset of ≈ 7 pN could be erroneously calculated to be 7.1 pN, for an error of 0.1 pN, which leads to a net contribution to the total work—when integrated over a distance of 100 nm—of $W_{\text{err}} = 10\text{pN}\cdot\text{nm} \approx 2.5 k_{\text{B}}T$. Given that we are primarily interested in dissipative work values which differ from one another by only a few $k_{\text{B}}T$, errors such as these can frustrate the analysis, resulting in artificially large work variances, and thus making it difficult to statistically differentiate energetic costs that only differ by a few $k_{\text{B}}T$.

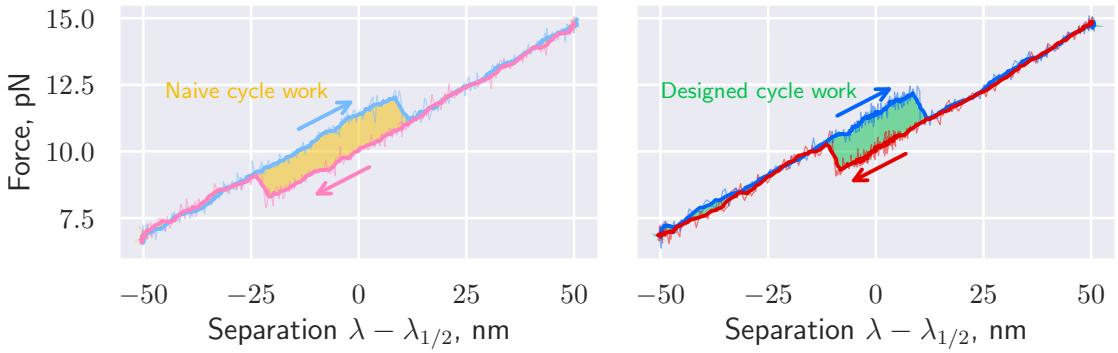

Figure 4.4: **Example force-separation curves showing the cycle work $W^{\mathrm{U}} + W^{\mathrm{F}}$ for naive (left) and designed (right) protocols.** The raw force-separation curve (thin) is smoothed by a Savitsky-Golay filter (thick). The area contained between the unfolding and refolding protocols visually represents the cycle work for the protocol pair.

Instead, we focus on calculating the *cycle work* (or hysteresis), which is obtained by adding the total work for forward and reverse protocols. Figure 4.4 shows two examples of cycle-work measurements for naive and designed protocols, with the shaded area between the force-separation curves depicting the cycle work.[2]

Within the linear-response approximation, the cycle work gives a direct measurement of twice the excess work

$$W_{\mathrm{cycle}} = W^{\mathrm{U}} + W^{\mathrm{R}} \tag{4.4a}$$

$$= W^{\mathrm{U}}_{\mathrm{ex}} + \Delta F^{\mathrm{U}} + W^{\mathrm{R}}_{\mathrm{ex}} + \Delta F^{\mathrm{R}} \tag{4.4b}$$

$$= 2W_{\mathrm{ex}} \ , \tag{4.4c}$$

where the final line follows from the fact that the equilibrium free energy difference for an unfolding protocol is the negative of the free energy difference for a folding protocol $\Delta F^{\mathrm{U}} = -\Delta F^{\mathrm{R}}$, and the excess work is predicted—within the linear-response approximation from Sec. 2.6.2—to be the same along a forward and reverse protocol: $W^{\mathrm{U}}_{\mathrm{ex}} = W^{\mathrm{R}}_{\mathrm{ex}} = W_{\mathrm{ex}}$.[3] If work measurements from unfolding and refolding protocols are taken from the same molecule, the free energy differences cancel exactly, effectively minimizing the effects of any experimental errors introduced in the force offset, and the resulting cycle work simply equals twice the excess work for the particular protocol.

[2]Figure 4.4 shows smoothed curves for the purpose of clarity, however the raw (unsmoothed) data are used when numerically calculating work values.

[3]There are other possible ways of combining work measurements to cancel out the free energy difference of the initial and final states, for instance the work difference $W_{\mathrm{naive}} - W_{\mathrm{designed}}$ in a particular direction, see App. B.2 for details.

The designed protocols are predicted by theory to reduce the excess work required for a fixed-duration control protocol, thus we expect to find that the average cycle work is reduced among designed protocols relative to their naive counterparts. Theory also suggests—as seen in Sec. 2.6.3—that such designed protocols are constructed so that the excess power, or rate of excess work accumulation, is constant along the protocol. Conversely, it predicts that the excess power along a naive protocol is maximized where the friction is largest.

4.2.1 Excess power in designed and naive protocols

To calculate the excess power in a particular region of control-parameter space $\Delta\lambda_{n+1,n} \equiv \{\lambda_n \leq \lambda_i < \lambda_{n+1}\}$, we calculate the work accumulated along each realization of a control protocol passing through this region as the sum

$$W^{(\Delta\lambda_{n+1,n})} = \sum_{\lambda_i=\lambda_n}^{\lambda_{n+1}} \frac{f(\lambda_{i+1}) + f(\lambda_i)}{2}(\lambda_{i+1} - \lambda_i) \ , \tag{4.5}$$

where the summand includes all extensions λ_i between the lower (λ_n) and upper (λ_{n+1}) bounds of the region $\Delta\lambda_{n+1,n}$. Without reference to any particular constraints on boundary conditions, the excess work in a particular interval $\Delta\lambda_{n+1,n}$ can be approximated by adding together the total work accumulated within a given region $\Delta\lambda_{n+1,n}$ from forward and reverse realizations of the same protocol class (designed or naive)

$$W_{\rm ex}^{(\Delta\lambda_{n+1,n})} \approx \frac{1}{2}\left(W_{\rm U}^{(\Delta\lambda_{n+1,n})} + W_{\rm R}^{(\Delta\lambda_{n+1,n})}\right) \ . \tag{4.6}$$

Here, because the excess work is predicted—within the linear-response regime—to be the same for unfolding and refolding protocols (4.4), the summand is simply twice the excess work (hence the factor of 1/2.)

Repeating across all protocol pairs—where again, the protocols in any given pair are from the same molecule—we find the average excess work $\langle W_{\rm ex}^{(\Delta\lambda_{n+1,n})}\rangle$ per interval. Thus, the average excess power in the region $\lambda \in [\lambda_n, \lambda_{n+1}]$ is the ratio

$$\left\langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \right\rangle = \frac{\langle W_{\rm ex}^{(\Delta\lambda_{n+1,n})}\rangle}{\Delta t^{(\Delta\lambda_{n+1,n})}} \ . \tag{4.7}$$

Here, $\Delta t^{(\Delta\lambda_{n+1,n})}$ is the amount of time that the protocol spends in the control-parameter range $\lambda \in [\lambda_n, \lambda_{n+1}]$. For equally sized regions $\Delta\lambda_{n+1,n}$, the time spent in any particular region is uniform for a naive protocol, but non-uniform for a designed protocol, where the time intervals are larger in regions near $\lambda_{1/2}$ and smaller in regions away from $\lambda_{1/2}$.

To directly compare the excess power across different protocol durations, we report a particular measure of excess power that is unchanged by variation in the protocol duration. Specifically, increasing the protocol duration manifests as decreasing total excess work, thus

we divide each excess power—whether it be from a naive or designed protocol—by the sum of average excess powers along an entire naive control protocol,

$$\left\langle \tilde{\mathcal{P}}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \right\rangle = \frac{\langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle}{\sum_{n=1}^{N} \langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive}} \ , \tag{4.8}$$

where $\langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive}$ is the average excess power along a naive protocol in the range $\lambda \in [\lambda_n, \lambda_{n+1}]$. This definition of the scaled excess power ensures that, along a naive protocol, the sum of scaled naive excess powers is normalized, $\sum_{n=1}^{N} \langle \tilde{\mathcal{P}}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive} = 1$ (the same need not be true, in general, for designed protocols). Because in the near-equilibrium regime [100] the excess power in either the numerator or denominator is predicted to scale the same with protocol duration ($\sim \tau^{-2}$), this scaled measure of excess power (4.8) in a particular control parameter region can be used to directly compare the excess power for different protocol classes, and across a range of protocol durations.

In practice, the denominator in (4.8) is numerically evaluated for equal bin sizes ($\Delta\lambda_{n+1,n}$) along a naive protocol by relating it to the excess work

$$\langle W_{\rm ex} \rangle_{\rm naive} = \sum_{n=1}^{N} \langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive} \Delta t^{(\Delta\lambda_{n+1,n})} \tag{4.9a}$$

$$= \frac{\tau}{N} \sum_{n=1}^{N} \langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive} \ , \tag{4.9b}$$

and thus the sum over excess powers is related to the average excess work as

$$\sum_{n=1}^{N} \langle \mathcal{P}_{\rm ex}^{(\Delta\lambda_{n+1,n})} \rangle_{\rm naive} = \frac{N \langle W_{\rm ex} \rangle_{\rm naive}}{\tau} \tag{4.10}$$

Figure 4.5 shows histograms of the scaled excess power (4.8) as a function of separation $\lambda - \lambda_{1/2}$, along naive and designed protocols for three different durations. As predicted by theory, naive protocols require a large amount of excess power near $\lambda_{1/2}$, where the generalized friction peaks, whereas the excess power along designed protocols is, by comparison, relatively uniform. This follows from the fact that the theory used to design minimum-dissipation protocols aims to make the excess power along designed protocols uniform as a function of separation (Sec. 2.6.3). For smaller protocol durations—and thus faster protocol speeds—the peak in the excess power of naive protocols broadens, and the excess power along the designed protocols becomes less uniform, indicating that the system is outside of the linear-response regime.

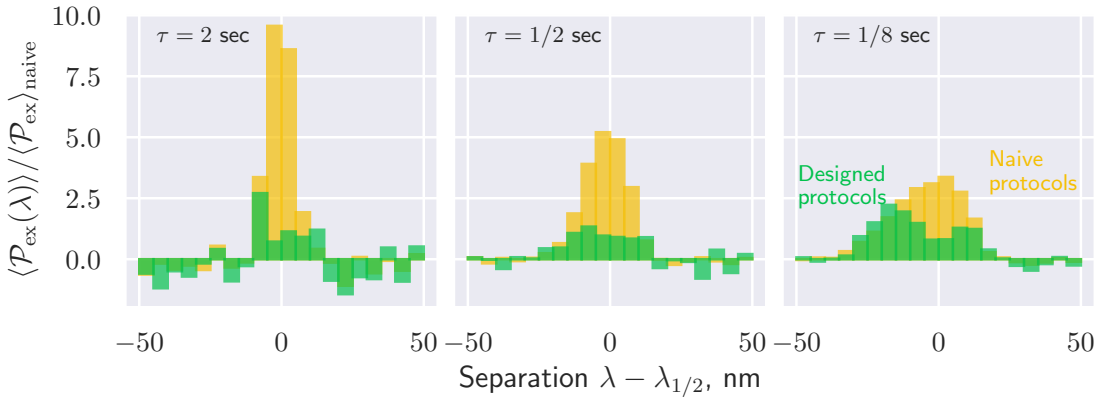

Figure 4.5: **Excess power as a function of separation for naive and designed protocols.** The excess power (normalized by the naive excess power) as a function of trap separation, for designed (green) and naive (yellow) protocols.

4.2.2 Cycle work and dissipation in designed and naive protocols

We obtain the average cycle work for each protocol class as a function of protocol duration by summing the total work measurements (4.3) (obtained by numerical integration) of selecting protocol pairs, each from the same molecule. Assuming linear response, the cycle work is twice the excess work (4.4) for that particular protocol class. Figure 4.6a shows the probability distributions for the cycle work along naive (yellow) and designed (green) protocols for three protocol durations. This shows that the work distributions for naive cycles are shifted to larger work values, relative to the designed-protocol work distributions, and that, as the distributions broaden for faster protocols (smaller protocol durations), the difference increases. Figure 4.6b shows the mean cycle work for designed and naive protocols as a function of protocol duration, for all 6 durations investigated experimentally. This shows that the designed protocols systematically require less excess work than their naive counterparts.

4.3 Protocol work ratios

The ratio between the excess work along a naive protocol and the excess work along a designed protocol quantifies the relative benefit of designing control protocols. Within the linear-response regime, theory predicts that the ratio of the average excess work along a designed protocol to the corresponding naive protocol is independent of the protocol duration τ and determined entirely by the friction coefficient [127]:

$$\frac{\langle W_{ex} \rangle_{naive}}{\langle W_{ex} \rangle_{designed}} = \frac{\overline{\zeta(\lambda_i)}}{\left(\overline{\zeta(\lambda_i)^{1/2}}\right)^2} , \tag{4.11}$$

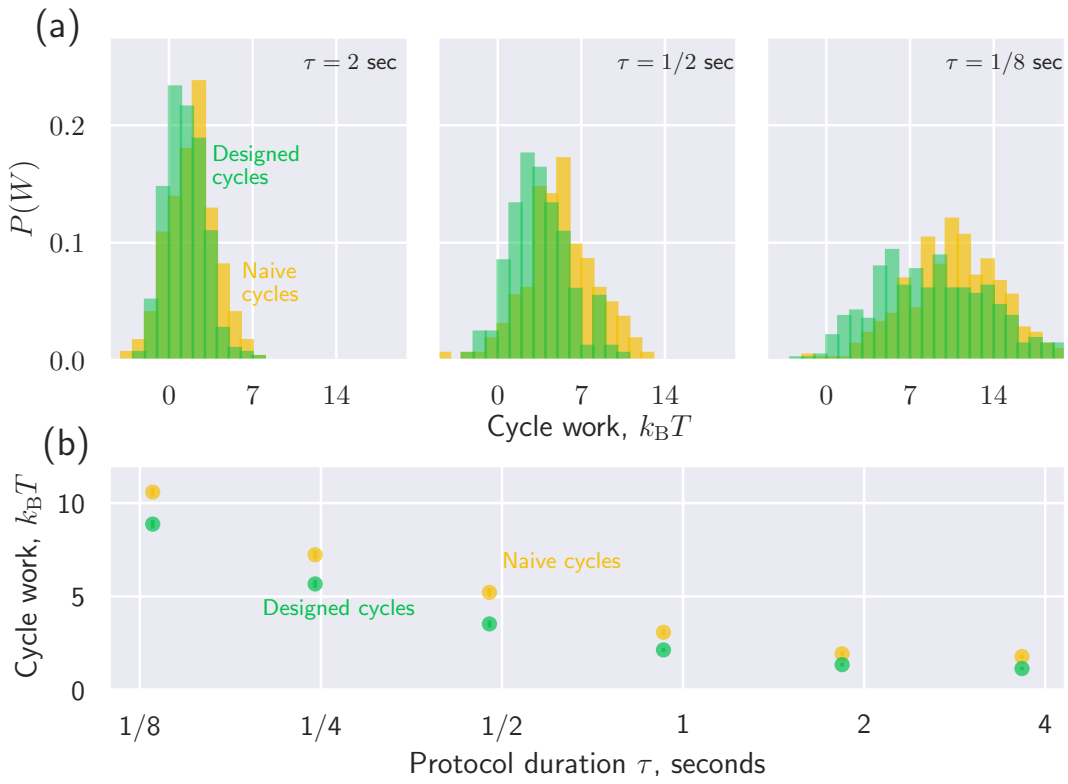

Figure 4.6: **(a) Cycle work for naive and designed protocols.** The distributions of cycle work $W^{\mathrm{U}}+W^{\mathrm{R}}$ for naive (yellow) and designed (green) protocols, for protocols ranging from slow (left) to fast (right). (b) Mean cycle work $\langle W^{\mathrm{U}} + W^{\mathrm{R}}\rangle$ during naive (yellow) and designed (green) protocols as a function of protocol duration.

where the overline $\overline{\cdots}$ indicates an average over λ, taken with respect to the uniform distribution. So $\overline{\zeta} = \frac{1}{\Delta\lambda_{\mathrm{tot}}}\int_{\lambda_{\mathrm{i}}}^{\lambda_{\mathrm{f}}}\zeta(\lambda)\mathrm{d}\lambda$, with $\Delta\lambda_{\mathrm{tot}} \equiv \lambda_{\mathrm{f}} - \lambda_{\mathrm{i}}$ the total displacement of a protocol.

In our measurements of equilibrium fluctuations of the hairpin (Ch. 3), we sampled more finely near $\lambda_{1/2}$ in order to resolve the features of the generalized friction more clearly. However, the derivation of (4.11) assumes that the bins are equally spaced, giving equal weighting to the generalized friction at all points, thus we need to generalize this result to unequal bin widths.

4.3.1 Excess work ratio for variable bin widths

Within the linear-response regime, the excess work can be calculated by the integral expression (2.69a). If the generalized friction is known at only N discrete values, the excess work can be approximated as

$$\langle W_{\mathrm{ex}}\rangle_{\Lambda} \approx \sum_{n=1}^{N}\zeta(\lambda_n)\left[\frac{\Delta\lambda_n}{\Delta t_n}\right]^2\Delta t_n \ , \tag{4.12}$$

where $\Delta\lambda_n$ and Δt_n are the distance traveled and time spent in bin n. After canceling factors of Δt_n, (4.12) simplifies to

$$\langle W_{\text{ex}}\rangle_\Lambda \approx \sum_{n=1}^{N} \Delta\lambda_n \left[\frac{\Delta\lambda_n}{\Delta t_n}\right] \zeta(\lambda_n) \ . \tag{4.13}$$

We next define the uniform spacing $\Delta\lambda_{\text{uni}} \equiv \Delta\lambda_{\text{tot}}/N$, and express the width of a non-uniform region as $\Delta\lambda_n = \Delta\lambda_{\text{uni}}\delta\lambda_n$. The weighting factors $\delta\lambda_n$ quantify the size of a particular bin relative to the uniform width; furthermore $\sum_{n=1}^{N}\delta\lambda_n = 1$ by definition. Using this, the excess work becomes

$$\langle W_{\text{ex}}\rangle_\Lambda \approx \frac{\Delta\lambda_{\text{tot}}}{N} \sum_{n=1}^{N} \left[\frac{\Delta\lambda_n}{\Delta t_n}\right] \zeta(\lambda_n)\delta\lambda_n \ . \tag{4.14}$$

For a naive (constant-velocity) protocol, $\Delta\lambda_n/\Delta t_n = \Delta\lambda_{\text{tot}}/\tau$, and the excess work is approximated as

$$\langle W_{\text{ex}}\rangle_{\text{naive}} \approx \frac{(\Delta\lambda_{\text{tot}})^2}{\tau N} \sum_{n=1}^{N} \zeta(\lambda_n)\delta\lambda_n \tag{4.15a}$$

$$\propto \overline{\zeta(\lambda_n)}_{\text{w}} \ , \tag{4.15b}$$

where the subscript 'w' on the overline now indicates that the generalized friction at each control parameter λ_n is weighted by the factor $\delta\lambda_n$.

Conversely, for a single control parameter, the designed protocol has a velocity that is proportional to the inverse square root of the generalized friction: $\Delta\lambda_n/\Delta t_n = A\zeta(\lambda_n)^{-1/2}$ (Sec. 2.6.3). Following the original derivation in [127], we fix the proportionality constant A by imposing the constraint of a fixed total duration:

$$\tau = \sum_{n=1}^{N} \Delta t_n \tag{4.16a}$$

$$= \sum_{n=1}^{N} \frac{\Delta\lambda_n}{A\zeta(\lambda_n)^{-1/2}} \tag{4.16b}$$

$$= \frac{\Delta\lambda_{\text{uni}}}{A} \sum_{n=1}^{N} \zeta(\lambda_n)^{1/2}\delta\lambda_n \tag{4.16c}$$

$$A = \frac{\Delta\lambda_{\text{tot}}}{N\tau}\overline{\zeta(\lambda_n)^{1/2}}_{\text{w}} \ . \tag{4.16d}$$

We substitute this expression for the proportionality constant A in (4.14) to show that near equilibrium the average excess work along a minimum-dissipation protocol is

$$\langle W_{\mathrm{ex}}\rangle_{\mathrm{designed}} \approx \frac{(\Delta\lambda_n)^2}{N\tau}\overline{\zeta(\lambda_n)_{\mathrm{w}}^{1/2}}\sum_{n=1}^{N}\zeta(\lambda_n)^{1/2}\delta\lambda_n \tag{4.17a}$$

$$= \frac{(\Delta\lambda_{\mathrm{tot}})^2}{N\tau}\left(\overline{\zeta(\lambda_n)_{\mathrm{w}}^{1/2}}\right)^2 . \tag{4.17b}$$

Thus, the ratio of (4.17b) and (4.15a) gives

$$\frac{\langle W_{\mathrm{ex}}\rangle_{\mathrm{naive}}}{\langle W_{\mathrm{ex}}\rangle_{\mathrm{designed}}} = \frac{\overline{\zeta(\lambda_n)_{\mathrm{w}}}}{\left(\overline{\zeta(\lambda_n)_{\mathrm{w}}^{1/2}}\right)^2} , \tag{4.18}$$

which takes the same form as the original result (4.11) with the replacement of the uniform spatial averages $\overline{\cdots}$ with the weighted averages $\overline{\cdots}_{\mathrm{w}}$.

4.3.2 Protocol-work ratio for DNA hairpins

For the DNA hairpin, the excess work ratio can be calculated by dividing the mean cycle work averaged over naive cycles by the cycle work averaged over designed cycles. The excess work ratio in (4.18) is expected to hold in the slow-protocol limit, where the hairpin remains in the near-equilibrium regime. Conversely, in the fast-protocol limit, the hairpin will not be able to relax throughout the course of the protocol, therefore never unfolding or refolding. In fact, in our discussion of alternate hairpin sequences in Chapter 5, we show that, for protocol durations $\tau < 1/2$ seconds, a significant fraction of the force extension curves for the slow-relaxing hairpin (3.1a) do not show reformation of the hairpin during the refolding protocols. In the absence of any force-rips, the force-separation curve of a particular realization is a monotonic function of separation, independent of the protocol class.[4] Thus, in this limit, both designed and naive protocols, on average, require the same amount of excess work, and thus the excess work ratio is unity.

Figure 4.7 shows the excess work ratio for the fast-relaxing hairpin as a function of protocol duration. For slow protocols (long durations), the experimentally obtained excess work ratio tends towards the long-duration excess work ratio (4.18), while for fast protocols (short durations), the experimental excess work ratios tend towards unity. For the slowest protocols ($\tau \approx 4$ sec), the experimental value obtained for the excess work ratio agrees (within uncertainty) with the theoretically predicted excess work ratio. These results suggest

[4]This can be seen, for instance, in Fig. 4.1. As the protocol duration decreases (and therefore the average protocol speed increases) the force-rips occur at later times in the protocol. Thus, for short enough durations, no force-rip events occur before the end of the protocol, and the entire force-separation curve is a monotonic function of separation.

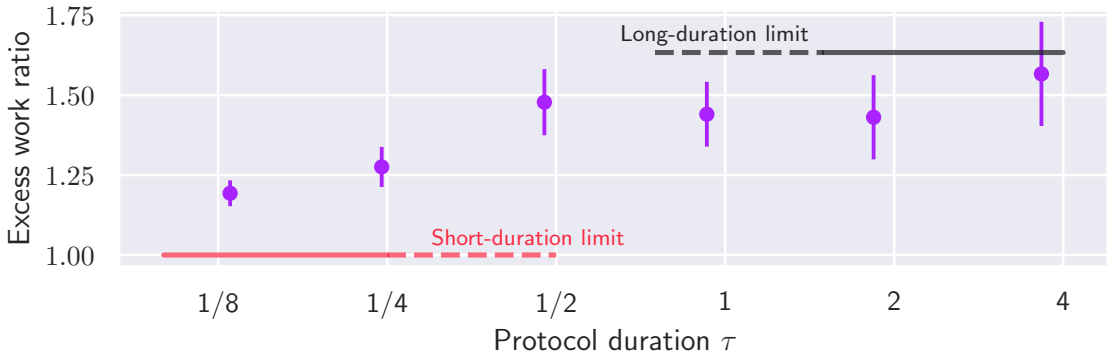

Figure 4.7: **Ratio of naive and designed protocol excess work for various protocol durations**. Experimentally obtained excess work ratios (purple points) as a function of protocol duration. For long durations, the experimental results tend toward the theoretical prediction (black line) while for short durations, they tend toward the theoretical expectation of unity (red line).

that, for even slower protocol durations ($\tau > 4$ sec), the excess work ratio may saturate the theoretically predicted value.

Chapter 5

DNA Hairpins III: robustness, variability, and conclusions

Chapters 3 and 4 presented results for a particular hairpin sequence (3.1b) that allowed relatively rapid folded-unfolded equilibration, such that transitions to the folded or unfolded conformations occurred even for the $\tau \approx 1/8$-second protocols. This feature allowed us to interrogate the hairpin's nonequilibrium response over a broad range of protocol durations. Here, we show how the qualitative and quantitative aspects of the analysis generalize to a different hairpin sequence (3.1a) and varying experimental conditions. In particular, we examine how the generalized friction (Fig. 3.3) is altered by changing the hairpin sequence, how the cycle work (Fig. 4.6) is affected, and to what extent these changes agree with theoretical predictions. Throughout this Chapter we refer to the hairpin sequence previously analyzed as the fast-relaxing hairpin, and the alternative sequence as the slow-relaxing hairpin.

5.1 Alternative hairpin sequence

The slow-relaxing hairpin sequence (3.1a) is a slight modification of the fast-relaxing sequence: it has the same stem length and loop size, but a different base-pair content in the stem. In comparison to the fast-relaxing hairpin, there are fewer GC base pairs near the terminal base pair of the slow-relaxing hairpin. This modification is expected to suppress the nucleation of the hairpin structure in the slow-relaxing hairpin relative to the fast-relaxing hairpin. As a result, the hopping dynamics are expected to be slower, leading to longer force relaxation times. Figure 5.1 compares the equilibrium fluctuations in the hopping regime for the slow- and fast-relaxing hairpin sequences. The dwell time of the hairpin to remain in either the folded or unfolded configuration is noticeably larger in the slow-relaxing hairpin (Fig. 5.1a), which is shown quantitatively by the relatively slow decay of the force autocovariance function in the slow-relaxing hairpin ($\tau_{\text{relax}} \approx 200$ ms), compared to the fast-relaxing hairpin ($\tau_{\text{relax}} \approx 30$ ms), as seen in Fig. 5.1c. By contrast, the

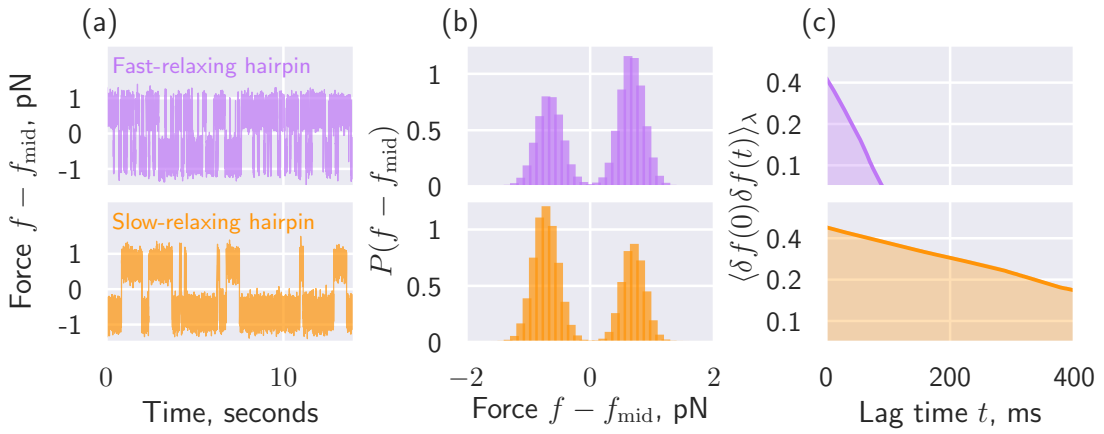

Figure 5.1: **Comparison of equilibrium analysis in the hopping regime for fast- and slow-relaxing hairpins.** (a) Sample force traces in the hopping regime for the fast-relaxing (top, purple) and slow-relaxing (bottom, orange) hairpin sequences. Raw forces f are reported relative to the distribution midpoint $f_{\mathrm{mid}} \equiv \frac{1}{2}(f_{\max} - f_{\min})$ to allow straightforward comparison between the two data sets. (b) Equilibrium force distributions are largely unchanged between the two sequences, while the force autocovariance function (c) decays much slower for the slow-relaxing hairpin.

equilibrium distribution of forces (and therefore force variance), as shown in Fig. 5.1b, is largely unchanged ($\langle \delta f^2 \rangle \approx 0.4$ pN^2 for both hairpin sequences).

We calculated the generalized friction, force variance, and integral relaxation time through an analogous process to that for the fast-relaxing hairpin (Ch. 3). Figure 5.2 compares the friction, force variance, and integral relaxation time of the slow-relaxing hairpin to the fast-relaxing hairpin (also shown in Fig. 3.3). The force variance in both hairpin sequences is similar, while the slow-relaxing hairpin has a much larger relaxation time, with a peak $\approx 100\times$ greater than the fast-relaxing hairpin. The generalized friction, as a result, also peaks at a value $\approx 100\times$ greater than in the fast-relaxing hairpin, suggesting that the near-equilibrium energy dissipation in DNA hairpin folding is highly sequence dependent.

The predicted minimum-dissipation control protocols in the fast and slow-relaxing hairpins are shown in Figure 5.2d,e for a 0.5 second protocol—which is the fastest protocol we were able to perform in the slow-relaxing hairpin sequence. Despite the large disparity in the relative magnitude of the generalized friction in the slow and fast-relaxing hairpin sequences, and since the designed protocol only depends on relative variation of the generalized friction coefficient, the designed control parameter velocity (Fig. 5.2d) and the designed protocols (Fig. 5.2e) are broadly similar for the two different hairpin sequences.

Figure 5.3 compares the distribution of cycle works for naive and designed protocols across several protocol durations, for both the fast-relaxing (top row) and slow-relaxing (bottom row) hairpin sequences. In both hairpins, the same qualitative trends hold for the distributions of cycle work: for shorter protocol durations the distributions widen and shift

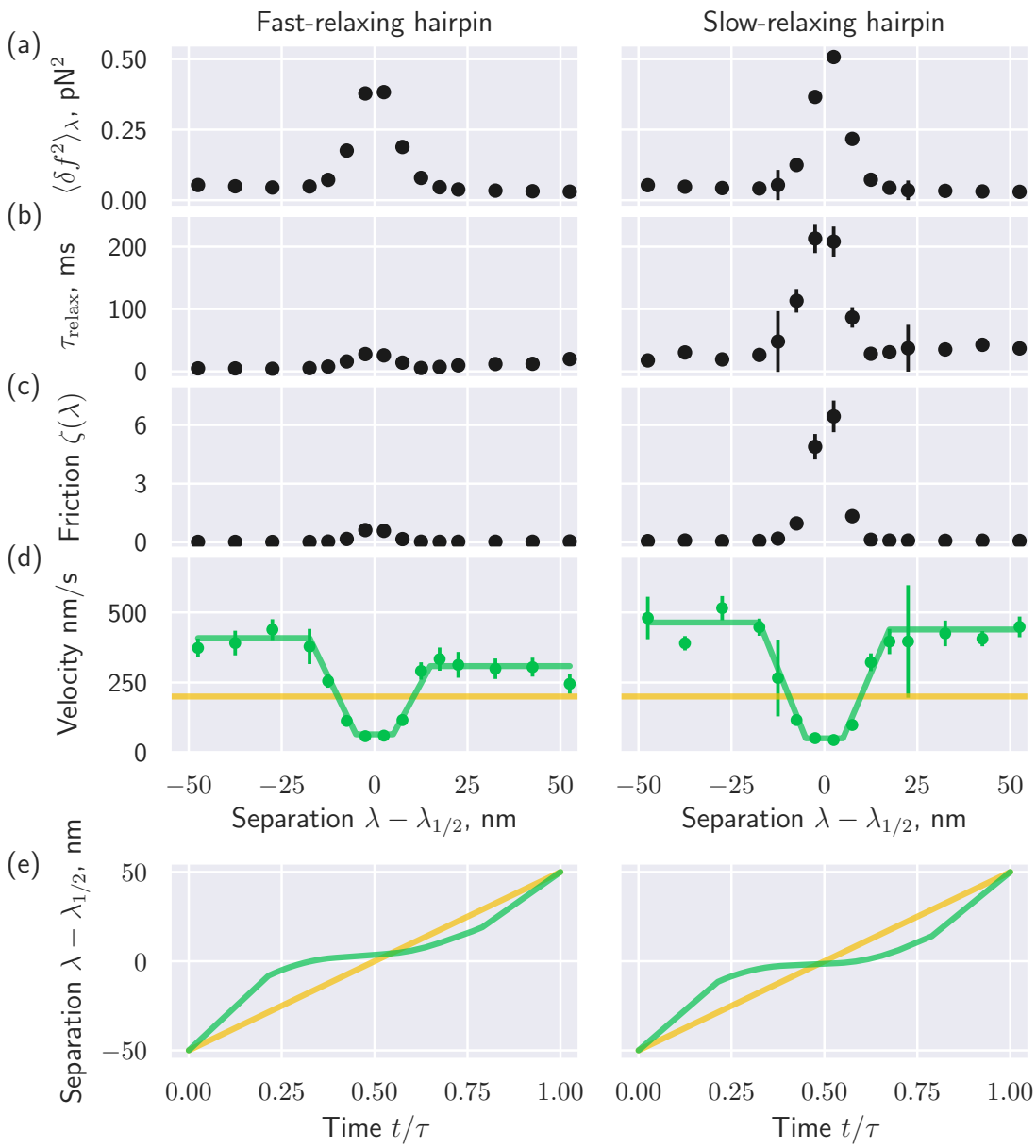

Figure 5.2: **Comparison of equilibrium statistics and designed protocols in fast- and slow-relaxing hairpins.** (a) Force variance as a function of trap separation for the fast-relaxing (left) and slow-relaxing (right) hairpin sequences are similar in magnitude and shape. (b) The relaxation time peak is $\approx 100\times$ higher for the slow-relaxing hairpin. (c) The generalized friction also peaks at a value $\approx 100\times$ greater in the slow-relaxing hairpin than the fast-relaxing hairpin. (d,e) Designed protocols for fast- and slow-relaxing hairpins appear qualitatively similar, but the slow-relaxing hairpin requires more extreme velocities.

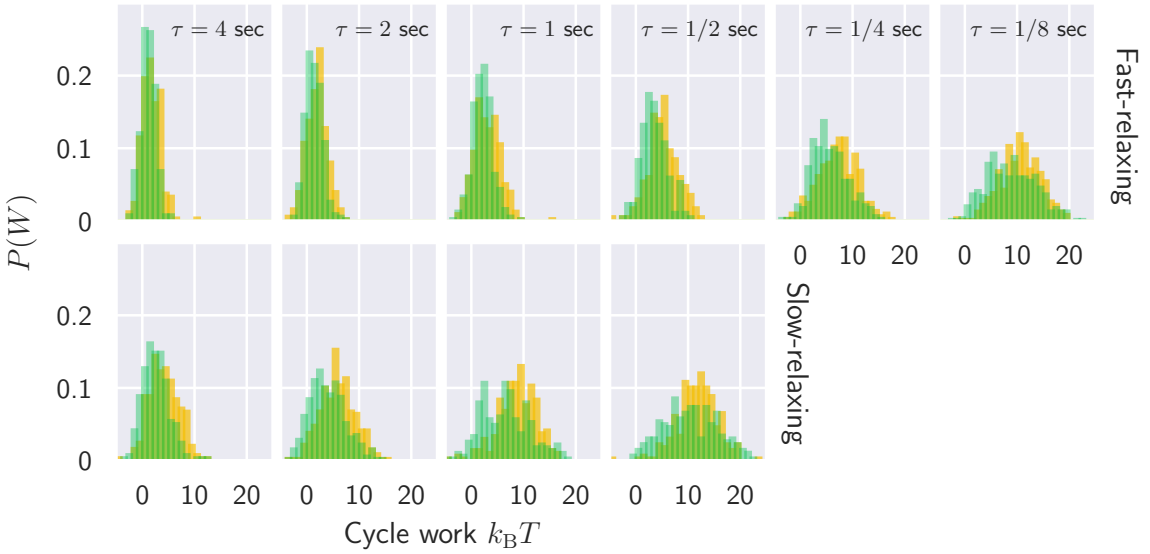

Figure 5.3: **Comparison of cycle work for fast and slow-relaxing hairpins.** The distributions of cycle work $W^{\mathrm{U}} + W^{\mathrm{R}}$ for naive (yellow) and designed (green) protocols for protocols ranging from slow (left) to fast (right), for both the fast-relaxing (top row) and slow-relaxing (bottom row) hairpin sequences. During the two shortest protocols ($\tau \approx 1/4, 1/8$ seconds), the dynamics of the slowly relaxing hairpin did not allow refolding, so this data was not analyzed.

towards higher dissipation. Furthermore, in both hairpin sequences, the distribution of cycle works for naive protocols is shifted higher relative to the designed protocols. For the fastest two protocols (with durations of $\tau = 1/4, 1/8$ seconds), a large fraction of the protocols in the slow-relaxing hairpin did not refold along the reverse protocols.

Finally, Fig. 5.4 compares the mean cycle work as a function of protocol duration in both the fast-relaxing (left) and slow-relaxing (right) hairpins. In both hairpin sequences, the designed protocols systematically require less excess work than the naive protocol at the same protocol duration. Furthermore, as predicted by theory, for a given protocol duration, the average cycle work is greater in the slow-relaxing hairpin than the fast-relaxing hairpin. Theoretically, the amplified peak friction values in the slow-relaxing hairpin increase the dissipation along a particular protocol.

5.2 Mean-variance trade-offs for excess work

Within the linear-response regime, the mean excess work is predicted to equal half the work variance [93]:

$$\beta \langle W_{\mathrm{ex}} \rangle_\Lambda \approx \frac{\beta^2}{2} \langle \delta W_{\mathrm{ex}}^2 \rangle_\Lambda \ . \tag{5.1}$$

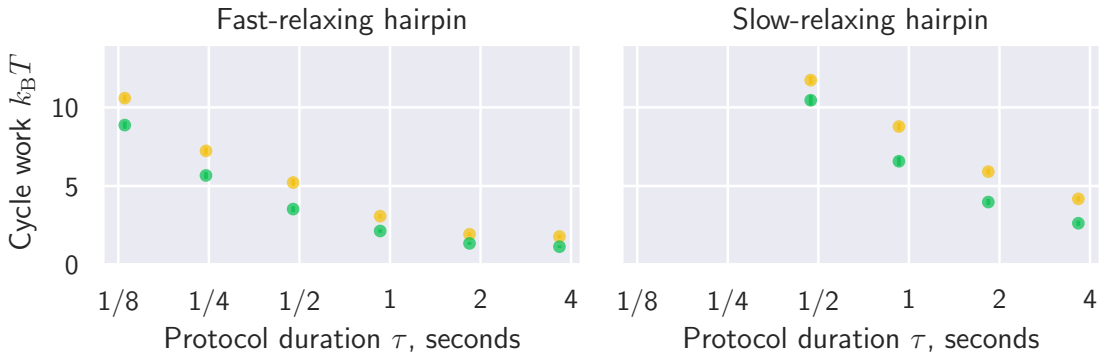

Figure 5.4: **Comparison of mean cycle work for fast- and slow-relaxing hairpins.** Mean cycle work as a function of protocol duration, during naive (yellow) and designed (green) protocols, for both fast-relaxing (left) and slow-relaxing (right) hairpin sequences.

This result follows from the fact that the the Jarzynski equality $\langle \exp(-\beta W_{\text{ex}}) \rangle = 1$ equates the (LHS) excess-work moment-generating function (2.6) to unity, and thus taking the natural logarithm requires that the cumulant-generating function is $\ln\langle \exp(-\beta W_{\text{ex}}) \rangle = 0$. Within the linear-response regime, work distributions are Gaussian and thus only the first two cumulants are nonzero, implying (5.1) [6, 93].

Figure 5.5 shows the mean cycle work—which is twice the excess work (assuming linear-response)—and half of the cycle-work variance as a function of protocol duration, for both naive and designed protocols, and for both the fast-relaxing and slow-relaxing hairpin sequences. The mean cycle work and half of the cycle-work variance coincide within experimental error, for both naive and designed protocols, for all but the fastest protocols in the fast-relaxing hairpin and for the slowest protocols in the slow-relaxing hairpin.

This is consistent with our expectation that at a given protocol duration, the slow-relaxing hairpin is driven further out of equilibrium than the fast-relaxing hairpin, and thus departs at a longer protocol duration from the linear-response equality of mean cycle work and half cycle-work variance. Moreover, when each of the two hairpins is driven far from equilibrium and the two quantities differ, the (half) variance of cycle work exceeds the mean work in the designed protocols, where the mean is intentionally reduced, independent of its effect on the work variance. This provides an example of a system in which protocols designed with the goal of minimizing mean work are different than those designed to minimize the work variance.

In fact, the trade-offs between minimizing mean work and work variance in nonequilibrium driving protocols were explored recently in the context of a 'breathing harmonic trap' by Solon and Horowitz [130], showing that the driving protocols that minimize mean work and work variance can be very different. Thus the strategies implemented to minimize the mean excess work represent one specific aspect in the more general scope of so-called

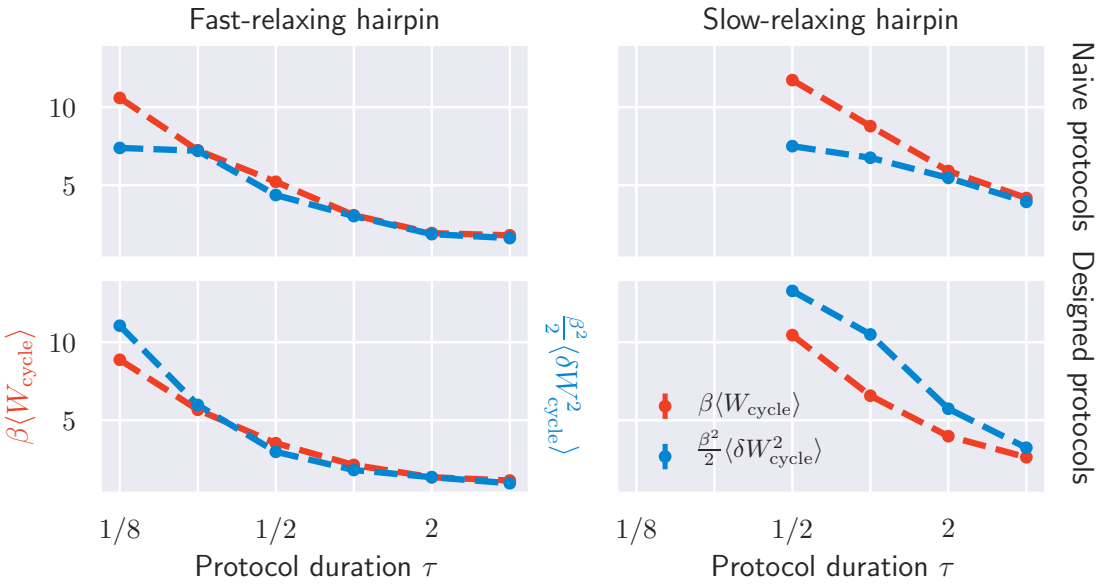

Figure 5.5: **Comparison between mean work and work variance for fast-relaxing and slow-relaxing hairpins.** Mean cycle work (red) and cycle-work variance (blue) as a function of protocol duration τ, for naive (top) and designed (bottom) protocols in the fast-relaxing (left) and slow-relaxing (right) hairpin sequences.

multi-objective optimization [165] that seeks to reconcile several (potentially competing) objectives at once, and thus the resulting optima need not bear any similarity to the result of single-objective optimization, such as minimizing the excess work.

5.3 Alternative buffer conditions

Finally, we investigated the effect of cations on the inferred friction coefficient. In particular, because DNA is a highly charged polyanion, the formation of secondary structure is facilitated by the presence of counterions to screen the electrical charge of the DNA backbone [166, 167, 168]. As a result, the thermodynamic stability of hairpin structures can vary greatly with the concentration of cations, with high counterion concentrations resulting in greater thermodynamic stability of secondary structures. Furthermore, divalent cations have a larger impact on the stabilization of secondary structures than monovalent cations do [169]. While counterions can stabilize the native hairpin structure, they may also lower the free energy of non-equilibrium secondary structures, thus increasing the probability of observing these kinetically trapped states. If the hopping behaviour of the DNA hairpin near $\lambda_{1/2}$ is dominated by electrostatic effects, then it is expected that the presence of a counterion could greatly alter the kinetics of loop formation. This could lead to a signif-

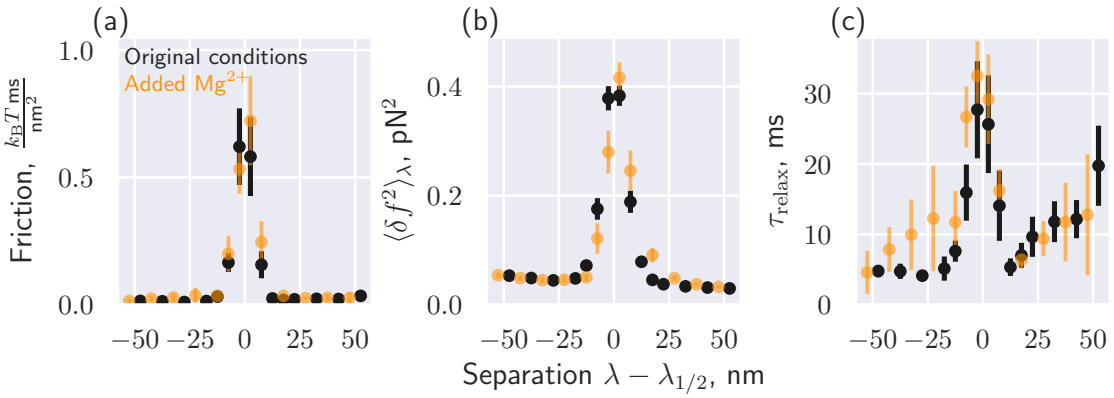

Figure 5.6: **Comparison of the force variance, integral relaxation time, and generalized friction of the fast-relaxing hairpin in two different buffers.** Friction (a), Force variance (b), and integral relaxation time (c) are largely unchanged between the solution used in the previous experiments in Chapters 3 and 4 (black) and after the addition of divalent Mg^{2+} ions (orange).

icant change in the hairpin relaxation time, and manifest with significant changes in the generalized friction (2.65).

In the experiments presented in Chapters 3 and 4, each hairpin was analyzed in 10 mM phosphate-buffered saline (PBS) with a 154 mM concentration of monovalent cations (NaCl). Here, we look at the effects on the generalized friction from adding divalent Mg^{2+}. In particular, we performed equilibrium sampling in a solution with 25 mM Tris-buffered saline (TBS),[1] 50 mM NaCl, and 5 mM $MgCl_2$.[2]

Figure 5.6 compares the generalized friction, force variance, and integral relaxation time of the fast-relaxing hairpin sequence (3.1b) in the conditions used for the analysis presented in the previous chapters, as well as after adding Mg^{2+} ions to the solution. Both experimental conditions produce nearly identical results, suggesting that the equilibrium dynamics are not significantly affected by adding divalent cations to the solution, at least for this hairpin sequence. The fast-relaxing hairpin sequence was designed to minimize the formation of alternative secondary structures, so changing ionic conditions may have more dramatic impact on other experimental systems, including different DNA hairpin sequences.

Furthermore, analysis of equilibrium quantities, such as the generalized friction, given no information on the formation of kinetically induced secondary structures. In particular, during a nonequilibrium unfolding/refolding protocol, the system may access kinetically favoured secondary structures that are not observed in the equilibrium statistics of the sys-

[1]Here, we use TBS in preference to PBS because phosphate buffers are known to react with divalent cations—such as Mg^{2+}—inhibiting their interactions with biomolecules [170].

[2]The presence of NaCl here is an inevitable contribution from the buffer, as explained in [169]. While the NaCl could, in principle, compete with the $MgCl_2$ cations, it was shown in [169] that this is not the case.

tem, but nonetheless play a significant role in the unfolding and refolding dynamics out of equilibrium. The presence of Mg^{2+} ions in solution may further stabilize these nonequilibrium structures, increasing the probability of the hairpin becoming trapped in one during the folding/unfolding process. A more detailed analysis of the role of buffer conditions and kinetically induced secondary-structure formation in systems like the DNA hairpin are beyond the scope of the present work, but represents a promising direction for future work. In fact, the importance of such kinetically favoured structures is predicted to play an important role in many biophysical processes, such as the formation of carboxysomes [171].

5.4 Discussion

In summary, we have sampled the equilibrium force fluctuations in DNA hairpins, displaying the dynamics of a two-state system in both the fast- and slow-relaxing sequences (Figs. 3.2, 5.1). We showed that the generalized friction coefficient—determined from such equilibrium fluctuations—can be used to design driving schedules (Figs. 3.3, 5.2) that significantly reduce the excess work compared with constant-velocity schedules (naive protocols) completed in the same total time (Fig. 4.6). This result holds for protocol durations that vary by a factor of ≈ 30, even when driven far from equilibrium (dissipating up to $\approx 10 k_B T$, which greatly exceeds the $\approx 1 k_B T$ energy fluctuations at equilibrium). These observations indicate that this near-equilibrium theory (Section 2.6.2) is still able to reduce dissipation even beyond the regime of the theory's strict validity.

This experiment represents the design (Chapter 3) and implementation (Chapter 4) of a single-molecule protocol that systematically reduces the nonequilibrium energy dissipation in a process constrained to finite duration.

These results have immediate applications in the streamlining of single-molecule experiments and steered molecular dynamics simulations [172]. For instance, when using the Jarzynski equality or Crooks fluctuation theorem to infer the free energy difference in a given process (such as protein unfolding), the farther the system is from equilibrium during experiment or simulation, the slower the rate of convergence and accuracy of the free energy estimator, which depends inversely on the energy dissipated [173]. Therefore, by sampling the equilibrium fluctuations of a biomolecular process, it should be possible to estimate the generalized friction coefficient across the control-parameter landscape; next, it would be possible to craft nonequilibrium protocols that dissipate significantly less energy, thereby speeding up the convergence and increasing the accuracy of any given free energy estimator.

There are tantalizing hints of molecular machines conserving energy while operating out of equilibrium: the $\varphi 29$ DNA packaging motor is more likely to slow down and pause at high packaging fractions, where the storing of additional DNA involves significantly higher dissipation [137, 174], and translating ribosomes facing RNA hairpins—that impose a large barrier to translation—change "gear," operating slower while crossing the barrier [43, 175].

Based on the theoretical framework presented here, both cases could be interpreted as examples in which the molecular machines implement driving protocols that proceed slower where the friction coefficient is higher, thereby reducing dissipation and increasing their efficiency. We hypothesize that a molecular biophysical system can waste less energy through naturally evolved dynamics that is rationalizable in terms of the generalized friction coefficient; specifically, such molecular motors may have evolved to slow down their operation in regions of their control-parameter space corresponding to high values of the friction coefficient as a way to harness fluctuations from the thermal bath, thus improving their operation efficiency.

The agreement of theory [109] and our experiments suggests extensions to more complex contexts. The rotary motor F_1-ATP synthase (Sec. 1.1.2) is known to be a remarkably efficient machine [176], where the F_o subunit—powered by proton flow down a concentration gradient—forces rotation of the γ-subunit, a molecular crankshaft that drives synthesis of ATP by F_1 [139]. After attaching a magnetic bead to the crankshaft of F_1 [177], one could—analogous to the procedure described in the previous chapters—use a magnetic tweezers to hold the bead at various angles so as to extract the equilibrium torque fluctuations of the rotary crankshaft, from which one could extract the friction coefficient at each position (in this experiment, the angle corresponds to the control parameter for driving F_1 in analogy to the optical trap separation for driving the unfolding of the DNA hairpin). One could then estimate the minimum-dissipation protocol and determine the ratio of energy input (work done to rotate the crankshaft) to energy output (ATP molecules synthesized) [177, 178] for designed and naive protocols. These ratios quantify the energetic efficiency with which the respective protocols induce F_1 to synthesize ATP, and their difference determines the energetic savings.

Part II

Dissipation in nonequilibrium systems through the lens of control theory

Chapter 6

Stochastic control in microscopic nonequilibrium systems

6.1 Introduction

Our original motivation for studying the minimal-dissipation control protocols in nonequilibrium systems was to gain a better understanding of the fundamental physics and operational constraints facing biological molecular machines. In Section 1.1 we saw two motivating examples of molecular machines—kinesin and ATP synthase—which play essential roles in the internal physics of biological cells. In light of the control-theoretical view of driven nonequilibrium systems outlined in Section 2.6, it is natural to look for ways in which the subcomponents of a given molecular machine interact with one another in a similar way. In particular, by viewing molecular machines as driven systems, we hope to apply the insights from control theory to molecular machines so as to better understand their dissipation, and move toward identifying the fundamental physical limitations that they face.

The experimental results presented in Chapters 3–5 provide a proof of principle that the linear-response theories used to derive the approximate form of the excess work in (2.69a) are useful in designing reduced-dissipation protocols in biomolecular systems. In fact, most theoretical efforts in this area have also focused on the deterministic protocols found in experiments, such as flipping or erasing a classical bit [117], or manipulating a biomacromolecule using optical traps or atomic force microscopy [127]. A deterministic protocol lends itself naturally to single-molecule experiments, where the same time-dependent driving protocol can be reliably repeated. Yet in biomolecular contexts, the nonequilibrium driving may be imposed by molecular machines that are themselves composed of strongly fluctuating protein components.

The F_oF_1 ATP synthase rotary motor serves as a paradigmatic model system, which uses rapid (presumably far-from-equilibrium) mechanical rotation of a crankshaft—itself driven by proton flow across a membrane—to drive synthesis of ATP molecules [41]. At ambient temperature, these soft-matter system components (such as the crankshaft of ATP

synthase) undergo strong conformational fluctuations, hence can only provide stochastic driving protocols to downstream systems (such as the F_1 subunit that synthesizes ATP). Thus, in order to probe the thermodynamics of stochastic driving in autonomous systems, we consider energetic costs that arise from a statistical distribution of control protocols—as opposed to a single, deterministic protocol.

In this chapter we generalize the linear-response formalism from [109] so that it quantifies energetic costs associated with statistical ensembles of control protocols. Our central result is that this variation in control protocols creates an additional energetic cost associated with slow operation, leading to work being minimized at finite protocol duration. Under the linear-response approximation, the lower bound on work (6.29) and optimal duration (6.28) take on simple forms. For a single control parameter operating within these limits with uniform friction coefficient and control-parameter velocity fluctuations, this implies an optimal mean driving velocity equal to the standard deviation of those stochastic velocity fluctuations (6.30a). Our theoretical formulation identifies the existence of a minimal cost for stochastic control – the only control modality available for living soft-matter systems.

6.2 Revisiting linear response

Now, we briefly revisit the linear-response approximations made in deriving the excess work in (2.69a), and in particular, how they affect the application of this theory to stochastic protocols. Throughout this chapter, our derivations make no explicit assumption about the nature of such stochasticity, which could imply an ensemble of individual protocols each of which is a deterministic function of time, but with randomly selected parameters—such as the periodic-potential ensemble considered in Sec. 6.4.1—or protocols which are generated by a stochastic equation of motion—as with the stochastic protocol ensembles considered in Sec. 6.4.2.

First, we recall that—within the linear-response regime—the excess power at time t in a given control protocol Λ, averaged over system responses, takes on the integral expression (by combining (2.61) and (2.62))

$$\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} = \dot{\lambda}^i(t) \int_{-\infty}^{t} \langle \delta f_j(0) \delta f_i(t-t') \rangle_{\boldsymbol{\lambda}(t)} \, \dot{\lambda}^j(t') \, \mathrm{d}t' \ . \tag{6.1}$$

In order to elucidate more precisely the applicability of the linear-response framework in Section 2.6.2, we derive the general conditions under which this excess power simplifies to $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} = \beta \dot{\lambda}^i \zeta_{ij}(\boldsymbol{\lambda}) \dot{\lambda}^j$ (2.68). For this simplification to be valid, the control-parameter velocity within the integrand of (6.1) must be approximated by its current value (at t)

$$\dot{\lambda}^j(t') \approx \dot{\lambda}^j(t) \ . \tag{6.2}$$

Taking this term outside of the integral in (6.1), and changing variables $t - t' \to t''$, (6.1) becomes the excess power discussed in Sec. 2.6.2, and introduced in [109].

Here we consider the next-order terms for $\dot{\lambda}^j(t)$ and derive conditions under which the Taylor series truncation in (6.2) is valid. (To consider the conditions under which the linear-response approximation is generally valid, it would also be necessary to consider higher-order response functions.)

Expanding the control-parameter velocity $\dot{\lambda}^j(t')$ in (6.1) about the time argument gives (to first-order in $t - t'$)

$$\langle \mathcal{P}_{\text{ex}} \rangle_{\Lambda(t)} \approx \dot{\lambda}^i(t) \int_{-\infty}^{t} \langle \delta f_j(0) \delta f_i(t - t') \rangle_{\boldsymbol{\lambda}(t)} \left[\dot{\lambda}^j(t) + \partial_{t'} \dot{\lambda}^j \Big|_t (t' - t) \right] \mathrm{d}t' \ . \tag{6.3}$$

This expansion is well approximated by (2.68) when

$$\dot{\lambda}^j(t) \gg (t' - t)\, \ddot{\lambda}^j(t) \tag{6.4}$$

for time separations $t' - t$ over which the conjugate-force autocovariance $\langle \delta f_i(0) \delta f_j(t) \rangle_{\boldsymbol{\lambda}(t)}$ is significantly greater than zero. Generally, (6.4) is satisfied when the protocol Λ is smooth and slowly varying over timescales less than the relaxation time of the conjugate forces. Furthermore, this smoothness constraint depends on the protocol velocity, but for deterministic protocols can always be satisfied as, in the long-duration limit, $\ddot{\lambda}^j$ becomes vanishingly small.

However, for control protocols generated by a stochastic equation of motion—so that the each individual protocol Λ is the solution to, for instance, a Langevin equation (1.4)—the validity of the smoothness constraint (6.4) in any particular limit is less clear. As mentioned in Section 2.2.3, the sample paths generated by an overdamped Langevin equation are continuous everywhere but differentiable nowhere, which means that the Taylor expansion of the control protocol velocity in (6.3) is unjustified—in fact, the control-protocol velocity itself is poorly defined in such a case.

However, the overdamped Langevin equation (1.4) is often obtained, through a limiting process, from an underdamped equation (2.32), where velocities are well defined. Furthermore, the approximation made in simplifying the underdamped Langevin equation to an overdamped Langevin equation is one of timescales: if the relaxation time of the velocity in an underdamped Langevin equation is significantly faster than any other relevant timescales in the problem, an overdamped approximation is justified. For instance, in observing a diffusing particle in a fluid, if the time between observations is shorter than the velocity relaxation time of the particle motion, an overdamped description will be consistent with observations [72].

Here, we are concerned with the smoothness of protocols Λ over timescales of the system response. Thus, the relevant comparison of timescales is between the relaxation times of the control parameter and the system. In particular, if the control parameter is sufficiently un-

derdamped, such that control-parameter velocities remain correlated over times which are long compared to the conjugate-force relaxation time of the system, then the protocol is effectively smooth over timescales relevant to (6.4), and the linear-response approximation in (6.2) is valid. For instance, a control parameter evolving according to an underdamped Langevin equation and confined with a harmonic potential of strength k_λ—so that its equation of motion is an underdamped Ornstein-Uhlenbeck process [84] (as in Sec. 6.4.2)—has velocity relaxation time $(\beta D_\lambda k_\lambda)^{-1}$, where D_λ is the diffusion coefficient of the control-parameter dynamics. Thus if the relaxation time of conjugate forces of the system is long compared with $(\beta D_\lambda k_\lambda)^{-1}$, then (6.4) is valid. Throughout this thesis, we refer to this constraint as the *locally smooth limit*, where the stochastic dynamics of a continuous control protocol are smooth on timescales relevant to the system response, even though they may still have large fluctuations over longer timescales.

In the context of the harmonic potential, a decrease in either the diffusion coefficient D_λ or the spring constant k_λ increases the velocity relaxation time and thus provides a mechanism to reach the locally smooth limit. In the deterministic theory, slowing down the control-parameter velocity results in better agreement with theory, whereas when the protocols are generated from a stochastic equation of motion there is an additional requirement that the control-parameter dynamics are sufficiently underdamped that (6.4) holds throughout the protocol.

6.3 Protocol ensembles and stochastic control

Now, instead of considering a single protocol $\Lambda : \boldsymbol{\lambda}_\mathrm{i} \to \boldsymbol{\lambda}_\mathrm{f}$, we consider an ensemble Ω of protocols, where each individual protocol Λ satisfies (6.4) and occurs with probability $p[\Lambda|\Omega]$. The excess power $\langle \mathcal{P}_\mathrm{ex} \rangle_{\Lambda(t)} \approx \dot{\lambda}^i \zeta_{ij}(\boldsymbol{\lambda}) \dot{\lambda}^j$ at time t during protocol $\Lambda \in \Omega$, averaged over system fluctuations, is now a random variable since $\dot{\lambda}^i$, $\zeta_{ij}(\boldsymbol{\lambda})$, and $\dot{\lambda}^j$ are all functions of the random protocol Λ. The excess power, averaged over system *and* protocol fluctuations, is

$$\langle \mathcal{P}_\mathrm{ex} \rangle_{\Omega(t)} \equiv \int \langle \mathcal{P}_\mathrm{ex} \rangle_{\Lambda(t)} \, p[\Lambda|\Omega] \, \mathcal{D}[\Lambda] \ , \tag{6.5}$$

where the integral is taken over all protocols and hence all instantaneous values of $\dot{\boldsymbol{\lambda}}(t)$. $\langle \cdots \rangle_{\Omega(t)}$ indicates an average over the instantaneous distribution of control-parameter positions and velocities at time t due to the protocol ensemble Ω. Furthermore, we assume—analogously to previous applications of the generalized friction framework—that the control protocol proceeds independently of the state of the system: the protocol receives no feedback from the system. In Chapter (9), we will explore the consequences of this assumption, in particular for its relevance to molecular machines.

6.3.1 Expansion of the excess power

By expanding the linear-response excess power $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)}$ in a Taylor series and assuming a weak-noise perturbation expansion of the protocols $\Lambda(t)$ about their mean value in the ensemble Ω at time t, we obtain a tractable expression for the protocol-ensemble average in (6.5). First, we Taylor expand the linear-response approximation of the average excess power about its mean [179]:

$$\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} = \sum_{n,m,l} \frac{1}{n!m!l!} \times \partial_{\dot\lambda^i}^n \partial_{\zeta_{ij}}^m \partial_{\dot\lambda^j}^l \langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} \Big|_{\langle \dot\lambda^i \rangle_{\Omega(t)}, \langle \zeta_{ij} \rangle_{\Omega(t)}, \langle \dot\lambda^j \rangle_{\Omega(t)}} (\delta \dot\lambda^i)^n (\delta \zeta_{ij})^m (\delta \dot\lambda^j)^l \ . \tag{6.6}$$

This expansion requires that $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)}$ is a smooth function of $\dot\lambda^i$, $\dot\lambda^j$, and ζ_{ij}, which clearly holds for $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} = \dot\lambda^i \zeta_{ij} \dot\lambda^j$. For notational simplicity, throughout this subsection, we will suppress the dependence of ζ_{ij} on the control parameter $\boldsymbol{\lambda}$.

Keeping all nonzero terms, the excess power is

$$\begin{aligned} \langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} = {} & \langle \dot\lambda^i \rangle_{\Omega(t)} \langle \zeta_{ij} \rangle_{\Omega(t)} \langle \dot\lambda^j \rangle_{\Omega(t)} + \langle \dot\lambda^i \rangle_{\Omega(t)} \langle \dot\lambda^j \rangle_{\Omega(t)} \delta\zeta_{ij} + \langle \dot\lambda^i \rangle_{\Omega(t)} \langle \zeta_{ij} \rangle_{\Omega(t)} \delta\dot\lambda^j \\ & + \langle \dot\lambda^j \rangle_{\Omega(t)} \langle \zeta_{ij} \rangle_{\Omega(t)} \delta\dot\lambda^i + \langle \dot\lambda^i \rangle_{\Omega(t)} \delta\zeta_{ij} \delta\dot\lambda^j + \langle \dot\lambda^j \rangle_{\Omega(t)} \delta\zeta_{ij} \delta\dot\lambda^i + \langle \zeta_{ij} \rangle_{\Omega(t)} \delta\dot\lambda^i \delta\dot\lambda^j \\ & + \delta\dot\lambda^i \delta\zeta_{ij} \delta\dot\lambda^j \ , \end{aligned} \tag{6.7}$$

because terms of fourth- and higher-order are trivially zero by the form of $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda_t}$. When we average $\langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)}$ over the protocol ensemble Ω,

$$\langle \mathcal{P}_{\rm ex} \rangle_{\Omega(t)} \equiv \int \langle \mathcal{P}_{\rm ex} \rangle_{\Lambda(t)} \, p[\Lambda(t)|\Omega] \, \mathcal{D}[\Lambda(t)] \ , \tag{6.8}$$

all terms vanish trivially which are linear in (protocol) fluctuations from the mean, leaving:

$$\begin{aligned} \langle \mathcal{P}_{\rm ex} \rangle_{\Omega(t)} = {} & \langle \dot\lambda^i \rangle_{\Omega(t)} \langle \zeta_{ij} \rangle_{\Omega(t)} \langle \dot\lambda^j \rangle_{\Omega(t)} + \langle \dot\lambda^i \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot\lambda^j \rangle_{\Omega(t)} + \langle \dot\lambda^j \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot\lambda^j \rangle_{\Omega(t)} \\ & + \langle \zeta_{ij} \rangle_{\Omega(t)} \langle \delta\dot\lambda^i \delta\dot\lambda^j \rangle_{\Omega(t)} + \langle \delta\dot\lambda^i \delta\zeta_{ij} \delta\dot\lambda^j \rangle_{\Omega(t)} \ . \end{aligned} \tag{6.9}$$

We assume the friction $\zeta_{ij}(\boldsymbol{\lambda})$ is a smooth function of the control parameter, which holds if all conjugate forces f_i are even under momentum reversal, except at a macroscopic phase transition. In the limit of weak noise [72], where the ensemble of protocols is tightly localized about its average, we expand the excess power perturbatively in noise strength. Specifically, we assume that the i,j-th component of the friction tensor evolves in accordance with the general linear stochastic differential equation (SDE):

$$\dot{\zeta_{ij}}(\boldsymbol{\lambda}, t) = a_{ij}(\boldsymbol{\zeta}, t) + \epsilon \, b_{ij}^{\ell}(\boldsymbol{\zeta}, t) \, \xi_{\ell}(t) \ , \tag{6.10}$$

where $\xi_\ell(t)$ is the ℓth element of a zero-mean (vector) white-noise process affecting control parameter λ^ℓ, with covariance $\langle \xi_\ell(t)\xi_m(t')\rangle = \delta_{\ell,m}\delta(t-t')$. $a_{ij}(\boldsymbol{\zeta}, t)$ is a function describing the deterministic behavior of ζ_{ij}, and $b^\ell_{ij}(\boldsymbol{\zeta}, t)$ is a third-rank tensor quantifying how fluctuations in each control parameter affect the friction tensor. The matching upper and lower indices on $b^\ell_{ij}(\boldsymbol{\zeta}, t)\xi_\ell(t)$ imply a sum over the index ℓ, accounting for the effects of all control-parameter fluctuations on the i, j-th component of the friction.

Following Gardiner [72], we make the small-noise perturbative expansion of $\zeta_{ij}(\boldsymbol{\lambda}, t)$ in the small parameter ϵ representing the magnitude of friction fluctuations:

$$\zeta_{ij}(\boldsymbol{\lambda}, t) = \zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t) + \epsilon\, \zeta^{(1)}_{ij}(\boldsymbol{\lambda}, t) + \epsilon^2\, \zeta^{(2)}_{ij}(\boldsymbol{\lambda}, t) + \cdots \ . \tag{6.11}$$

Here $\zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t)$ is the solution to the deterministic equation $\mathrm{d}_t\zeta_{ij} = a_{ij}(\boldsymbol{\zeta}, t)$ (hence independent of fluctuations in $\boldsymbol{\lambda}$), whereas each of the $\zeta^{(n)}_{ij}(\boldsymbol{\lambda}, t), n > 0$, has stochastic contributions. $\epsilon\zeta^{(1)}_{ij}(\boldsymbol{\lambda}, t)$ is the leading-order correction to the deterministic solution in the limit of weak noise. Expanding $a_{ij}(\boldsymbol{\zeta}, t)$ in ϵ—similarly to (6.11)—and grouping terms with a common power of ϵ, yields the first-order correction to the deterministic approximation in the weak-noise limit, given by the solution to

$$\dot{\zeta}^{(1)}_{ij}(\boldsymbol{\lambda}, t) = -u\left(\zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t)\right) + b^\ell_{ij}\left(\zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t)\right)\xi_\ell(t) \ . \tag{6.12}$$

Here $u\left(\zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t)\right) \equiv -\partial a\left(\zeta^{(0)}_{ij}\right)/\partial\zeta^{(0)}_{ij}$. We have used the initial condition that $\zeta^{(1)}_{ij}(\boldsymbol{\lambda} = \boldsymbol{\lambda}_0, t = 0) = 0$, which is equivalent to all protocols starting at the same point in control-parameter space. (6.12) is simply a time-dependent Ornstein-Uhlenbeck process [72].

Thus the two factors involving ζ_{ij} appearing in (6.9) are, to leading-order in ϵ,

$$\langle \zeta_{ij}\rangle_{\Omega(t)} = \zeta^{(0)}_{ij}(\boldsymbol{\lambda}, t) = \zeta_{ij}(\langle\boldsymbol{\lambda}\rangle_{\Omega(t)}, t) \tag{6.13a}$$

$$\delta\zeta_{ij} = \sum_{m=1}^{\infty} \epsilon^m \zeta^{(m)}_{ij} \approx \epsilon\zeta^{(1)}_{ij}(\boldsymbol{\lambda}, t) \ . \tag{6.13b}$$

(6.13a) shows that, to order ϵ in the weak-noise limit, the average friction along the protocol ensemble is the friction along the average protocol. (6.13b) shows that, to lowest order, the fluctuations in the friction can be approximated by the solution to a time-dependent Ornstein-Uhlenbeck process (6.12).

We now perform a similar analysis of the control-parameter velocity dynamics. To begin, we assume that—similar to the generalized friction ζ_{ij} (6.10)—the dynamics are described by a linear SDE,

$$\dot{\lambda}^i = a^i_\lambda(\lambda^i, t) + \epsilon b^i_\lambda(\lambda^i, t)\xi_i(t) \ , \tag{6.14}$$

where $\xi_i(t)$ is the component of the (vector) white-noise process which affects λ^i.

The average control-parameter velocity is

$$\langle \dot{\lambda}^i \rangle_{\Omega(t)} = \langle a^i_\lambda(\lambda^i, t) \rangle_{\Omega(t)} \ , \tag{6.15}$$

where the average is taken over protocol fluctuations at a given time t within the protocol ensemble Ω. Fluctuations in the control-parameter velocity are then

$$\delta\dot{\lambda}^i \equiv \dot{\lambda}^i - \langle a^i_\lambda(\lambda^i, t) \rangle_{\Omega(t)} \tag{6.16a}$$

$$= a^i_\lambda(\lambda^i, t) - \langle a^i_\lambda(\lambda^i, t) \rangle_{\Omega(t)} + \epsilon b^i_\lambda(\lambda^i, t)\xi_i(t) \ . \tag{6.16b}$$

Again, following Gardiner [72] we expand a^i_λ in a small parameter ϵ,

$$a^i_\lambda(\lambda^i, t) \approx a^{i,(0)}_\lambda + \epsilon a^{i,(1)}_\lambda$$

$$= \langle a^i_\lambda(\lambda^i, t) \rangle_{\Omega(t)} + \epsilon a^{i,(1)}_\lambda \ , \tag{6.17a}$$

from which it follows that $a^i_\lambda(\lambda^i, t) - \langle a^i_\lambda(\lambda^i, t) \rangle_{\Omega(t)} \approx \epsilon a^{i,(1)}_\lambda$, where $a^{i,(1)}_\lambda$ is the first-order correction to the dynamics of control parameter λ^i, analogous to $\zeta^{(1)}_{ij}$ in (6.11). Fluctuations in the control-parameter velocity are thus (to order ϵ) $\delta\dot{\lambda}^i \approx \epsilon \left[a^{i,(1)}_\lambda + b^i_\lambda(\lambda^i, t)\xi_i(t) \right]$.

With this weak-noise approximation in terms of mean and linear-order fluctuations in $\zeta_{ij}(\boldsymbol{\lambda})$, $\dot{\lambda}^i$, and $\dot{\lambda}^j$, the covariance terms in the excess power expansion (6.9) are

$$\langle \dot{\lambda}^i \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \dot{\lambda}^i \rangle_{\Omega(t)} \left\langle \zeta^{(1)}_{ij} a^{j,(1)}_\lambda \right\rangle_{\Omega(t)} + \langle \dot{\lambda}^i \rangle_{\Omega(t)} \left\langle \zeta^{(1)}_{ij} b^j_\lambda \xi_j(t) \right\rangle_{\Omega(t)} \right] \ , \tag{6.18a}$$

$$\langle \dot{\lambda}^j \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \dot{\lambda}^j \rangle_{\Omega(t)} \left\langle \zeta^{(1)}_{ij} a^{i,(1)}_\lambda \right\rangle_{\Omega(t)} + \langle \dot{\lambda}^j \rangle_{\Omega(t)} \left\langle \zeta^{(1)}_{ij} b^i_\lambda \xi_i(t) \right\rangle_{\Omega(t)} \right] \ , \tag{6.18b}$$

$$\langle \zeta_{ij} \rangle_{\Omega(t)} \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \zeta_{ij} \rangle_{\Omega(t)} \left\langle a^{i,(1)}_\lambda a^{j,(1)}_\lambda \right\rangle_{\Omega(t)} + \langle \zeta_{ij} \rangle_{\Omega(t)} \left\langle b^i_\lambda b^j_\lambda \xi_i(t) \xi_j(t) \right\rangle_{\Omega(t)} \right] \ , \tag{6.18c}$$

$$\langle \delta\dot{\lambda}^i \delta\zeta_{ij} \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \mathcal{O}(\epsilon^3) \ , \tag{6.18d}$$

where for notational simplicity $b^i_\lambda(\lambda^i, t) \equiv b^i_\lambda$, and the dependence of $\zeta^{(1)}_{ij}$ on $\boldsymbol{\lambda}, t$ is suppressed. We henceforth neglect the final $\mathcal{O}(\epsilon^3)$ term.

In the excess power expansion (6.9) the friction-control parameter velocity covariance terms (6.18a,6.18b) are negligible compared to the control-parameter velocity covariance term (6.18c) when

$$\frac{\langle \dot{\lambda}^j \rangle_{\Omega(t)}}{\langle \zeta_{ij} \rangle_{\Omega(t)}} \ll \frac{\left\langle a^{i,(1)}_\lambda a^{j,(1)}_\lambda \right\rangle_{\Omega(t)} + \left\langle b^i_\lambda b^j_\lambda \xi_i(t) \xi_j(t) \right\rangle_{\Omega(t)}}{\left\langle \zeta^{(1)}_{ij} a^{i,(1)}_\lambda \right\rangle_{\Omega(t)} + \left\langle \zeta^{(1)}_{ij} b^i_\lambda \xi_i(t) \right\rangle_{\Omega(t)}} \ . \tag{6.19}$$

In this limit, substituting (6.13a) into (6.18), the expansion of the excess power (6.6) simplifies to

$$\langle \mathcal{P}_{\text{ex}} \rangle_{\Omega(t)} \approx \langle \dot{\lambda}^i \rangle_{\Omega(t)} \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \dot{\lambda}^j \rangle_{\Omega(t)} + \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} \ . \tag{6.20}$$

This quantifies the excess work associated with completing an ensemble Ω of protocols, in terms of the average protocol $\langle \Lambda \rangle_\Omega$ (defined by the path taken by $\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}$) and the control-parameter velocity covariance $\langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)}$ along that path.

To explore the limits in which (6.19) holds, we derive an explicit expression for the friction fluctuations $\delta\zeta_{ij}$, given by the solution to the time-dependent Ornstein-Uhlenbeck process from (6.12):

$$\zeta_{ij}^{(1)}(\boldsymbol{\lambda}, t) = \int_0^t b_{ij}^\ell \left(\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, t') \right) e^{-\int_{t'}^t u\left(\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, s)\right) \mathrm{d}s} \mathrm{d}\xi_k(t') \ . \tag{6.21}$$

Here we have imposed the initial condition that all protocols begin at the same point in control-parameter space, equivalent to $\zeta_{ij}^{(1)}(\boldsymbol{\lambda}, t)$ vanishing at the start of the protocol, so the boundary term at $t = 0$ in (6.21) vanishes [72].

We now consider the case where, throughout the protocol, the trajectories of $\zeta_{ij}(\boldsymbol{\lambda}, t)$ are at steady state in the reference frame which is comoving with the deterministic solution $\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, t)$. Here, the fluctuations about the deterministic value of $\zeta_{ij}(\boldsymbol{\lambda}, t)$ are independent of time. This is the same constraint placed on the stochastic protocols in the ensemble considered in Sec. 6.4.2. $u\left(\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, t)\right)$ represents a time-dependent variation in the first-order correction about the deterministic limit of (6.10), so this constraint requires that $u\left(\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, t)\right) = 0$. Hence at steady state the exponential term in the integral expression (6.21) becomes unity.

If we also assume that the diffusion tensor for the ζ_{ij} dynamics is a constant $b_{ij}^\ell \left(\zeta_{ij}^{(0)}(\boldsymbol{\lambda}, t') \right) = b_{ij}^{\ell,(0)}$, then $\zeta_{ij}^{(1)}(\boldsymbol{\lambda}, t)$ is independent of $\boldsymbol{\lambda}$, and the integral expression simplifies greatly:

$$\zeta_{ij}^{(1)}(t) \approx b_{ij}^{\ell,(0)} \int_0^t \mathrm{d}\xi_\ell(t) \tag{6.22a}$$

$$= b_{ij}^{\ell,(0)} \xi_\ell(t) \ , \tag{6.22b}$$

where the final equality follows again from the initial condition that all protocols begin from the same point, and we use the Ito convention to evaluate the stochastic integral [72]. Subject to these assumptions, we write (6.18) in terms of the parameters of the weak-noise

expansion,

$$\langle \dot{\lambda}^i \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \dot{\lambda}^i \rangle_{\Omega(t)} \left\langle a_\lambda^{j,(1)} b_{ij}^{\ell,(0)} \xi_\ell(t) \right\rangle_{\Omega(t)} + \langle \dot{\lambda}^i \rangle_{\Omega(t)} \left\langle b_\lambda^j \xi_j(t) b_{ij}^{\ell,(0)} \xi_\ell(t) \right\rangle_{\Omega(t)} \right] \tag{6.23a}$$

$$\langle \dot{\lambda}^j \rangle_{\Omega(t)} \langle \delta\zeta_{ij} \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \dot{\lambda}^j \rangle_{\Omega(t)} \left\langle a_\lambda^{i,(1)} b_{ij}^{\ell,(0)} \xi_\ell(t) \right\rangle_{\Omega(t)} + \langle \dot{\lambda}^j \rangle_{\Omega(t)} \left\langle b_\lambda^i \xi_i(t) b_{ij}^{\ell,(0)} \xi_\ell(t) \right\rangle_{\Omega(t)} \right] \tag{6.23b}$$

$$\langle \zeta_{ij} \rangle_{\Omega(t)} \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} = \epsilon^2 \left[\langle \zeta_{ij} \rangle_{\Omega(t)} \left\langle a_\lambda^{i,(1)} a_\lambda^{j,(1)} \right\rangle_{\Omega(t)} + \langle \zeta_{ij} \rangle_{\Omega(t)} \left\langle b_\lambda^i b_\lambda^j \xi_i(t) \xi_j(t) \right\rangle_{\Omega(t)} \right] . \tag{6.23c}$$

Substituting (6.23) and the white-noise property $\langle \xi_i(t) \xi_j(t) \rangle = \delta_{ij}$ into (6.19) gives

$$\frac{\langle \dot{\lambda}^j \rangle_{\Omega(t)}}{\langle \zeta_{ij} \rangle_{\Omega(t)}} \ll \frac{\left\langle a_\lambda^{i,(1)} a_\lambda^{j,(1)} \right\rangle_{\Omega(t)} + \left\langle \left(b_\lambda^i\right)^2 \right\rangle_{\Omega(t)}}{\left\langle a_\lambda^{i,(1)} b_{ij}^{\ell,(0)} \xi_\ell(t) \right\rangle_{\Omega(t)} + \left\langle b_\lambda^i b_{ij}^{i,(0)} \right\rangle_{\Omega(t)}} . \tag{6.24}$$

If this holds for all i, j at all points in the average protocol, then the friction-velocity covariance terms in (6.9) can be neglected. The inequality holds trivially in the asymptotic limit $a_\lambda^{i,(1)} \to 0$ of long protocol durations (slow average control-parameter velocities), where (6.24) becomes

$$\frac{\langle \dot{\lambda}^j \rangle_{\Omega(t)}}{\langle \zeta_{ij} \rangle_{\Omega(t)}} \ll \frac{\left\langle \left(b_\lambda^i\right)^2 \right\rangle_{\Omega(t)}}{\left\langle b_\lambda^i b_{ij}^{i,(1)} \right\rangle_{\Omega(t)}} , \tag{6.25}$$

and the LHS becomes arbitrarily small as $\langle \dot{\lambda}^i \rangle_{\Omega(t)} \to 0$.

Thus, when the ensemble of control protocols is tightly localized around the average protocol—such that the friction varies little over control-parameter values with significant support—the excess power (averaged over protocol and system fluctuations) is well approximated by expanding $\langle \mathcal{P}_{\text{ex}} \rangle_{\Lambda(t)}$ (2.68) about the mean values of its arguments $\dot{\lambda}^i$, $\dot{\lambda}^j$, and $\zeta_{ij}(\boldsymbol{\lambda})$ to give (6.20).

Time integration of (6.20) gives the average excess work required to perform a random protocol sampled from Ω,

$$\langle W_{\text{ex}} \rangle_\Omega = \int_0^\tau \left[\langle \dot{\lambda}^i \rangle_{\Omega(t)} \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \dot{\lambda}^j \rangle_{\Omega(t)} + \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} \right] \mathrm{d}t , \tag{6.26}$$

where $\langle \cdots \rangle_\Omega$ indicates an average over all protocols $\Lambda \in \Omega$ weighted by $p[\Lambda|\Omega]$, and the upper integration bound τ is the duration of the average protocol $\langle \Lambda \rangle$. The first RHS term resembles (2.69a), quantifying the cost associated with fast operation, while the second term quantifies the energetic cost resulting from variability in the protocol velocities. Both terms are integrated along the (deterministic) average protocol specified by the average velocity

$\langle \dot{\boldsymbol{\lambda}} \rangle_{\Omega(t)}$. Thus, in the weak protocol-noise limit the effect of variable control only depends on the friction along this average path and the variation in velocities as a function of time.

6.3.2 Lower bound on excess work

The Cauchy-Schwarz inequality gives a lower bound for the first RHS term in (6.26) involving the thermodynamic length $\mathcal{L}(\langle \Lambda \rangle_\Omega)$ (2.71) between the initial and final states of the average protocol $\langle \Lambda \rangle_\Omega$ [124]. This leads to a lower bound on the excess work achieved at a finite protocol duration τ,

$$\langle W_{\rm ex} \rangle_\Omega \geq \frac{\mathcal{L}(\langle \Lambda \rangle_\Omega)^2}{\tau} + \left\langle \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} \right\rangle_\Omega \tau \ , \tag{6.27}$$

where we write the average of an instantaneous quantity over the protocol ensemble as $\langle \cdots \rangle_\Omega \equiv \tau^{-1} \int_0^\tau \cdots \, {\rm d}t$.

This lower bound represents a trade-off between the first RHS term quantifying the energetic costs associated with pushing a system out of equilibrium (scaling as τ^{-1} with protocol duration) and the second term quantifying the average contribution of protocol fluctuations to excess work, which increases with τ if $\zeta_{ij}(\boldsymbol{\lambda})$ is positive definite (assumed in what follows).

If the control-parameter velocity variance is independent of the average velocity, this lower bound is minimized at a finite protocol duration

$$\tau^{\rm opt} = \frac{\mathcal{L}(\langle \Lambda \rangle_\Omega)}{\left\langle \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} \right\rangle_\Omega^{1/2}} \ , \tag{6.28}$$

revealing a fundamental lower bound on the excess work,

$$\langle W_{\rm ex} \rangle_\Omega \geq 2 \left\langle \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta\dot{\lambda}^i \delta\dot{\lambda}^j \rangle_{\Omega(t)} \right\rangle_\Omega^{1/2} \mathcal{L}(\langle \Lambda \rangle_\Omega) \ . \tag{6.29}$$

This lower bound is saturated when the average protocol $\langle \Lambda \rangle_\Omega$ follows the geodesic from $\boldsymbol{\lambda}_{\rm i}$ to $\boldsymbol{\lambda}_{\rm f}$ (similar to the deterministic case). The existence of a lower bound on work realized at finite protocol duration constitutes the main result of this chapter. In the linear-response regime, the lower bound and optimal protocol duration take on the simple forms in (6.29) and (6.28), respectively.

For a single control parameter, (6.29) and (6.28) can be recast solely in terms of intensive quantities as a lower bound on the average excess force $\langle f_{\rm ex} \rangle_\Omega \equiv \langle W_{\rm ex} \rangle_\Omega / \Delta\lambda$ produced by an optimal mean control-parameter velocity $\langle \dot{\lambda} \rangle_\Omega^{\rm opt} \equiv \Delta\lambda / \tau^{\rm opt}$. When the friction and control-parameter velocity variance are both uniform, the excess force bound and optimal velocity

simplify to

$$\langle \dot{\lambda} \rangle_{\Omega}^{\text{opt}} = \sqrt{\langle \delta \dot{\lambda}^2 \rangle_{\Omega}} \tag{6.30a}$$

$$\langle f_{\text{ex}} \rangle_{\Omega} \geq 2\zeta \langle \dot{\lambda} \rangle_{\Omega}^{\text{opt}} \ . \tag{6.30b}$$

The optimal mean velocity is the root-mean-squared control-parameter velocity fluctuations, producing a mean excess force equal to twice the Stokes drag on the control parameter when moving at the optimal mean velocity through the 'viscous' control-parameter space subject to generalized friction coefficient ζ.

In the specific case where—across the entire protocol—the integral relaxation time is constant and equals $\tau^{\text{R}} = (\beta \langle \zeta_{ij}(\langle \boldsymbol{\lambda} \rangle_{\Omega(t)}) \langle \delta \dot{\lambda}^i \delta \dot{\lambda}^j \rangle_{\Omega(t)} \rangle_{\Omega})^{-1}$, our lower bound (6.29) reduces to Machta's bound on entropy production of a stochastically driven process [180]. This equality is achieved in the one-dimensional drift-diffusion process considered by Machta when protocol fluctuations come from the interaction of the control parameter with a thermal reservoir at the same temperature as the reservoir producing system fluctuations. Thus our derived lower bound (6.29) generalizes Machta's bound to systems with variable integral relaxation times and arbitrary fluctuations of the control parameter. Appendix C.1 provides a more comprehensive discussion.

6.4 Model ensembles

We illustrate our theoretical approximation (6.27) using two model protocol ensembles. In each case, the system is a Brownian particle with unit mass evolving according to an overdamped Langevin equation on a one-dimensional potential. Driving forces are produced by a harmonic potential $E(x, \lambda) = \frac{1}{2} k_{\text{trap}} [x - \lambda(t)]^2$ (Sec. 2.7.1), with trap strength k and control parameter $\lambda(t)$ the time-dependent potential minimum (Fig. 6.1). To saturate the excess work bounds in (6.27,6.29), we restrict our attention to protocol ensembles where the average protocol $\langle \Lambda \rangle_{\Omega}$ is the minimum-work protocol [109].

In each case, the system is a Brownian particle evolving according to an overdamped Langevin equation (1.4),

$$\dot{x} = -\beta D \partial_x E(x, \lambda) + \sqrt{2D} \xi(t) \ , \tag{6.31}$$

where x is the particle position, $-\partial_x E(x, \lambda)$ is the force experienced by the particle due to the potential $E(x, \lambda)$, D is the diffusion coefficient, $\beta \equiv (k_{\text{B}} T)^{-1}$ is the inverse temperature of the heat bath, and $\xi(t)$ is a zero mean white-noise process. The details of numerical implementation of this SDE are outlined in App. A.2.2.

For one control parameter, the theoretical minimum excess work for an ensemble Ω of driving protocols operating within the linear-response regime (6.27) with a constant control-

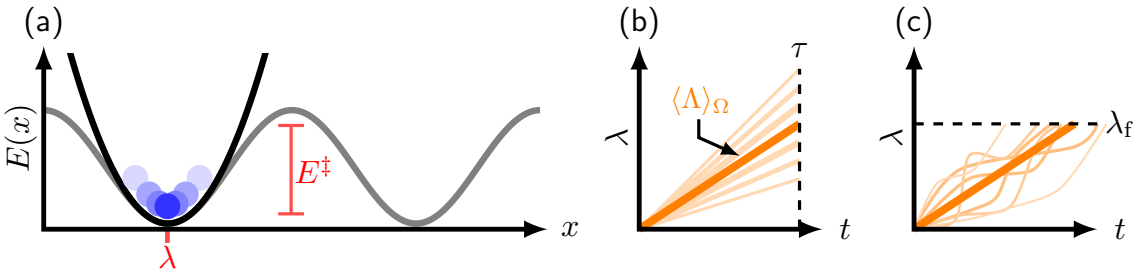

Figure 6.1: **Schematic of model system and protocol ensembles.** (a) Brownian particle diffusing in a harmonic potential, with harmonic trap minimum the control parameter, over an underlying periodic potential with barrier height $E^\ddagger$ between adjacent energetic minima. (b,c) Protocol samples from the zero-barrier periodic potential ensemble and the stochastically driven ensemble (for both cases $E^\ddagger = 0$). Bold lines denote average protocols.

parameter velocity variance simplifies to

$$\langle W_{\rm ex}\rangle_\Omega \geq \frac{\mathcal{L}(\langle\Lambda\rangle_\Omega)^2}{\tau} + \langle\zeta(\lambda)\rangle_\Omega\langle\delta\dot{\lambda}^2\rangle_\Omega\,\tau \ . \tag{6.32}$$

6.4.1 Periodic-potential ensemble

Here, the harmonic trap is driven over an underlying periodic potential $E_{\rm period}(x) = -\frac{1}{2}E^\ddagger\cos\pi x$ with energy barrier $E^\ddagger$ between adjacent wells (Fig. 6.1a). Thus the total potential experienced by the particle is

$$E(x,\lambda) = \frac{1}{2}k_{\rm trap}\left[x-\lambda(t)\right]^2 - \frac{1}{2}E^\ddagger\cos(\pi x) \ , \tag{6.33}$$

which is the model potential discussed in 2.7.2 (for $\ell = 2$). Here, the control parameter $\lambda(t)$ is the time-dependent minimum of the harmonic potential, and $E^\ddagger$ is the energy barrier separating consecutive minima of the underlying potential.

The generalized friction $\zeta(\lambda)$ (2.65) is nonuniform over the control-parameter landscape whenever $\beta E^\ddagger \neq 0$, and can be calculated with the one-dimensional form of (2.67) [118, 126]:

$$\zeta(\lambda) = \frac{1}{\beta D}\int_{-\infty}^{\infty}\frac{\left[\partial_\lambda\Pi_{\rm eq}(x|\lambda)\right]^2}{\pi(x|\lambda)}{\rm d}x \ , \tag{6.34}$$

for equilibrium cumulative distribution function $\Pi_{\rm eq}(x|\lambda) \equiv \int_{-\infty}^{x}\pi(x'|\lambda){\rm d}x'$ and system diffusion coefficient D.

Figure 6.2 shows—for various barrier heights $\beta E^\ddagger$—the generalized friction as a function of the control parameter λ, and the corresponding minimum-work protocols. Because the minimum-work protocols proceed with a control-parameter velocity $\dot{\lambda} \propto \zeta(\lambda)^{-1/2}$ (see Sec. 2.6.3), for non-zero barriers $\beta E^\ddagger \neq 0$ the protocols proceed slower near $\lambda = \ell/2$, where the system is near the peak of the underlying periodic potential, and the total po-

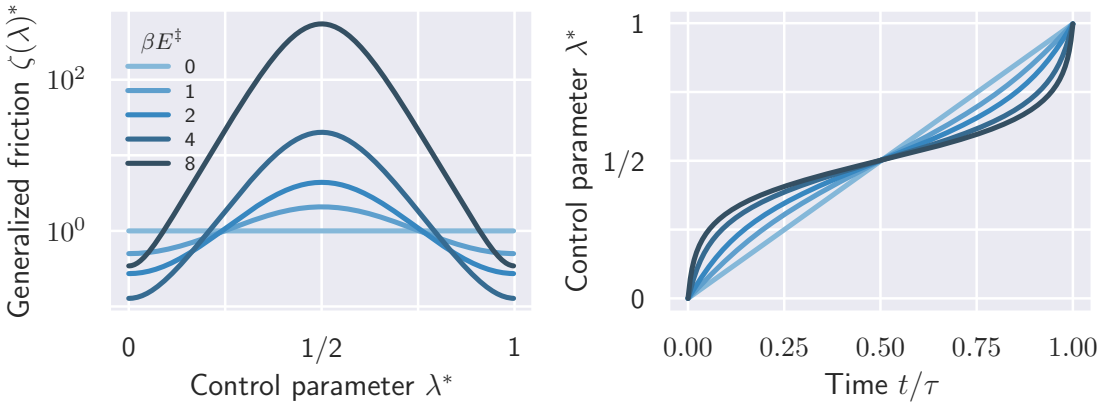

Figure 6.2: **Generalized friction and minimum-work protocols for various periodic barriers.** (a) The generalized friction $\zeta(\lambda)^* = \zeta(\lambda)/\gamma$ is nondimensionalized by the zero-barrier limit γ, and the friction is periodic over one full period ℓ of the potential. (b) Corresponding minimum-work protocols for $\lambda^* = \lambda/\ell$ as a function of t/τ.

tential (6.33) is bistable. As a result, the optimal driving protocols in this system resemble those considered in Chapters 3-5 for a DNA hairpin (see, for instance, Fig. 3.4b).

We examine a protocol ensemble where each protocol Λ completes the minimum-work path with an average velocity $\langle\dot{\lambda}\rangle_\Lambda$ randomly sampled from a Gaussian distribution with mean $\langle\dot{\lambda}\rangle_\Omega$ and variance $\langle\delta\dot{\lambda}^2\rangle_\Omega$. Each protocol has instantaneous velocity $\dot{\lambda} \propto [\zeta(\lambda)]^{-1/2}$ with the proportionality fixed by the prescribed average velocity $\langle\dot{\lambda}\rangle_\Lambda$. The ensemble-mean control-parameter velocity $\langle\dot{\lambda}\rangle_\Omega = \langle\Delta\lambda\rangle_\Omega/\tau$ is chosen so that the average protocol $\langle\Lambda\rangle_\Omega$ completes the control parameter change $\langle\Delta\lambda\rangle_\Omega$ in a prescribed time τ. Thus, the protocol ensemble Ω consists of protocols Λ which all take a fixed duration τ, but any given realization will travel a variable total distance. For each protocol with a particular average velocity $\langle\dot{\lambda}\rangle_\Lambda$, the system is initialized in the corresponding periodic steady-state for the minimum-work protocol. The periodic steady state is achieved when

$$p(x, t + \tau_\mathrm{L}) = p(x, t) \ , \tag{6.35}$$

where τ_L is the time it takes for the harmonic trap to traverse one periodic image of the underlying potential.

Exact solution in the zero-barrier limit ($\beta E^\ddagger = 0$)

In general, exact solutions for the excess work are intractable for $\beta E^\ddagger \neq 0$. However in the zero-barrier limit, where $\beta E^\ddagger = 0$, the exact solution for the protocol-ensemble averaged excess work coincides with the theoretical prediction (6.32). In this limit, the generalized friction is independent of the control-parameter value, and thus the minimum-work protocol

proceeds at a constant velocity $\dot{\lambda}(t) = \dot{\lambda}$ along any given realization (see Fig. 6.2). Direct integration of (6.31) for a constant control-parameter velocity gives

$$x(t) = x_0 e^{-\beta D k t} + \int_0^t e^{-\beta D k (t-s)} \left[\beta D k \lambda(t) + \sqrt{2D}\,\xi(s)\right] \mathrm{d}s \ . \tag{6.36}$$

This has mean position

$$\langle x \rangle_{\Lambda(t)} = \lambda(0) e^{-\beta D k t} - \int_0^t e^{-\beta D k (t-s)} \beta D k \dot{\lambda}\, s \ \mathrm{d}s \ . \tag{6.37}$$

Integration by parts gives

$$\langle x \rangle_{\Lambda(t)} = \lambda(t) - \int_0^t e^{-\beta D k (t-s)} \dot{\lambda}\, \mathrm{d}s \tag{6.38a}$$

$$= \lambda(t) - \frac{\dot{\lambda}}{\beta D k} \left(1 - e^{-\beta D k t}\right) \ . \tag{6.38b}$$

For long times ($t \to \infty$), the mean asymptotes to

$$\lim_{t\to\infty} \langle x \rangle_{\Lambda(t)} = \lambda(t) - \frac{\dot{\lambda}}{\beta D k} \ , \tag{6.39}$$

which lags the trap minimum by a distance $\dot{\lambda}/\beta D k$ [131].

The fluctuating work accumulated for a particular realization of the stochastic process of duration τ is (2.35)

$$W \equiv \int_0^\tau \partial_{t'} E(x, t')\, \mathrm{d}t' \tag{6.40a}$$

$$= -\int_0^\tau k \dot{\lambda} \left[x(t') - \lambda(t')\right] \mathrm{d}t' \ , \tag{6.40b}$$

for potential $E(x, t')$ taken from (6.33) (for $\beta E^\ddagger = 0$). The average work accumulated for a protocol of duration τ is

$$\langle W \rangle_\Lambda = -\int_0^\tau k \dot{\lambda} [\langle x \rangle_{\Lambda_{t'}} - \lambda(t')] \mathrm{d}t' \ . \tag{6.41}$$

Given that, for the potential considered here, there is no free energy change along any protocol, the total work equals the excess work.

Substituting for $\langle x \rangle_{\Lambda_{t'}}$ using (6.38b) gives the exact mean excess work [131]

$$\langle W_{\rm ex} \rangle_\Lambda = k\dot\lambda^2 \int_0^\tau \left[\int_0^{t'} e^{-\beta D k s} \mathrm{d}s \right] \mathrm{d}t' \tag{6.42a}$$

$$= \frac{\dot\lambda^2}{\beta D} \int_0^\tau \left[1 - e^{-\beta D k t'} \right] \mathrm{d}t' \tag{6.42b}$$

$$= \frac{\dot\lambda^2}{\beta D} \left[\tau - \frac{1 - e^{-\beta D k \tau}}{\beta D k} \right] . \tag{6.42c}$$

This differs from the linear-response approximation $\langle W_{\rm ex}^{\rm LR} \rangle_\Lambda = \dot\lambda^2 \tau / (\beta D)$ (which for this model is (6.32)) through the second term, representing the exponential decay of an initial transient.

Thus for a system which is initially in the nonequilibrium steady state (NESS) for the protocol velocity $\dot\lambda$, the exact mean excess work for a protocol duration of τ is [131]

$$\langle W_{\rm ex} \rangle_\Lambda = \frac{\dot\lambda^2}{\beta D} \tau \, . \tag{6.43}$$

We now generalize this result to randomly choose the (constant) control-parameter velocity at the start of each protocol. We consider the average of the fluctuating excess work as a path integral over all possible constant-velocity protocols,

$$\langle W_{\rm ex} \rangle_\Omega = \int p\left(\lambda_{[0,\tau]}, x_{[0,\tau]}\right) W_{\rm ex}\left(\lambda_{[0,\tau]}, x_{[0,\tau]}\right) \mathcal{D}\left[\lambda_{[0,\tau]}\right] \mathcal{D}\left[x_{[0,\tau]}\right] \tag{6.44a}$$

$$= \int p\left(\lambda_{[0,\tau]}\right) \mathcal{D}\left[\lambda_{[0,\tau]}\right] \int p(x_{[0,\tau]} \mid \lambda_{[0,\tau]}) W_{\rm ex}\left(\lambda_{[0,\tau]}, x_{[0,\tau]}\right) \mathcal{D}\left[x_{[0,\tau]}\right] \, . \tag{6.44b}$$

Here the subscript $[0,\tau]$ indicates that the variable represents a particular realization of the random process on the time interval $[0,\tau]$, and therefore the integral is taken over all paths of the processes $x_{[0,\tau]}$ and $\lambda_{[0,\tau]}$. The nested integral is the mean excess work for a single protocol Λ,

$$\langle W_{\rm ex} \rangle_\Lambda = \int p\left(x_{[0,\tau]} | \lambda_{[0,\tau]}\right) W_{\rm ex}\left(x_{[0,\tau]}, \lambda_{[0,\tau]}\right) \mathcal{D}\left[x_{[0,\tau]}\right] \, . \tag{6.45}$$

Thus, the mean work, averaged over all paths $x_{[0,\tau]}$ of the system variable as well as all protocols $\lambda_{[0,\tau]}$, is

$$\langle W_{\rm ex} \rangle_\Omega = \int p\left(\lambda_{[0,\tau]}\right) \langle W_{\rm ex} \rangle_\Lambda \mathcal{D}\left[\lambda_{[0,\tau]}\right] \, . \tag{6.46}$$

For the ensemble of constant-velocity protocols starting at equilibrium for the initial control-parameter value, the excess work $\langle W_{\text{ex}} \rangle_\Lambda$ is (6.42c), and the ensemble average is

$$\langle W_{\text{ex}} \rangle_\Omega = \int p\left(\lambda_{[0,\tau]}\right) \frac{\dot{\lambda}^2}{\beta D} \left[\tau - \frac{1 - e^{-\beta D k \tau}}{\beta D k}\right] \mathcal{D}\left[\lambda_{[0,\tau]}\right] \tag{6.47a}$$

$$= \frac{\langle \dot{\lambda}^2 \rangle_\Omega}{\beta D} \left[\tau - \frac{1 - e^{-\beta D k \tau}}{\beta D k}\right] . \tag{6.47b}$$

When the ensemble is initialized in the corresponding NESS, the excess work is (6.43), which when averaged over the protocol ensemble gives

$$\langle W_{\text{ex}} \rangle_\Omega = \frac{\langle \dot{\lambda}^2 \rangle_\Omega}{\beta D} \tau \ . \tag{6.48}$$

Both (6.47b) and (6.48) depend on $p(\lambda_{[0,\tau]})$ only through $\dot{\lambda}$, so in both cases, the same mean excess work is produced for any ensemble with constant velocities chosen from any distribution with a given mean and variance.

In the zero-barrier limit, the generalized friction is constant and equal to the viscous friction coefficient $\zeta = \gamma = 1/\beta D$. Thus, the thermodynamic length of the minimum work protocol is $\mathcal{L}(\langle \Lambda \rangle_\Omega) = \sqrt{\zeta} \langle \dot{\lambda} \rangle_\Omega \tau$ (2.71), and by using the identity $\langle \dot{\lambda}^2 \rangle_\Omega = \langle \dot{\lambda} \rangle_\Omega^2 + \langle \delta \dot{\lambda}^2 \rangle_\Omega$, the protocol ensemble produces an exact mean excess work

$$\langle W_{\text{ex}} \rangle_\Omega = \zeta \langle \dot{\lambda}^2 \rangle_\Omega \tau = \frac{\mathcal{L}(\langle \Lambda \rangle_\Omega)^2}{\tau} + \zeta \langle \delta \dot{\lambda}^2 \rangle_\Omega \tau \ , \tag{6.49}$$

which equals the linear-response approximation in (6.27).

Numerical results for non-zero barriers ($\beta E^\ddagger \neq 0$)

Figure 6.3 shows a comparison of numerical calculations to theoretical predictions, for several average protocol distances and periodic barrier heights. For no underlying barrier ($\beta E^\ddagger = 0$), the numerical data agrees exactly with the analytical solution. $\beta E^\ddagger = 1$ shows good agreement, but with increasing barrier height, the linear-response approximation in (6.4) begins to break down for rapid protocols, and the numerical results depart from the theoretical predictions. However, for all barriers explored, even those for which (6.4) does not hold, the excess work is minimized at finite protocol duration. Appendix C.2 further discusses the departure of theoretical predictions from numerical results in this model system.

6.4.2 Stochastically driven protocols

Here, the protocol itself evolves according to a dynamic stochastic process, transitioning between given initial and final control-parameter values λ_i and λ_f in a variable duration τ' (see Fig. 6.1c). Unlike the fixed-duration boundary conditions on the protocols considered

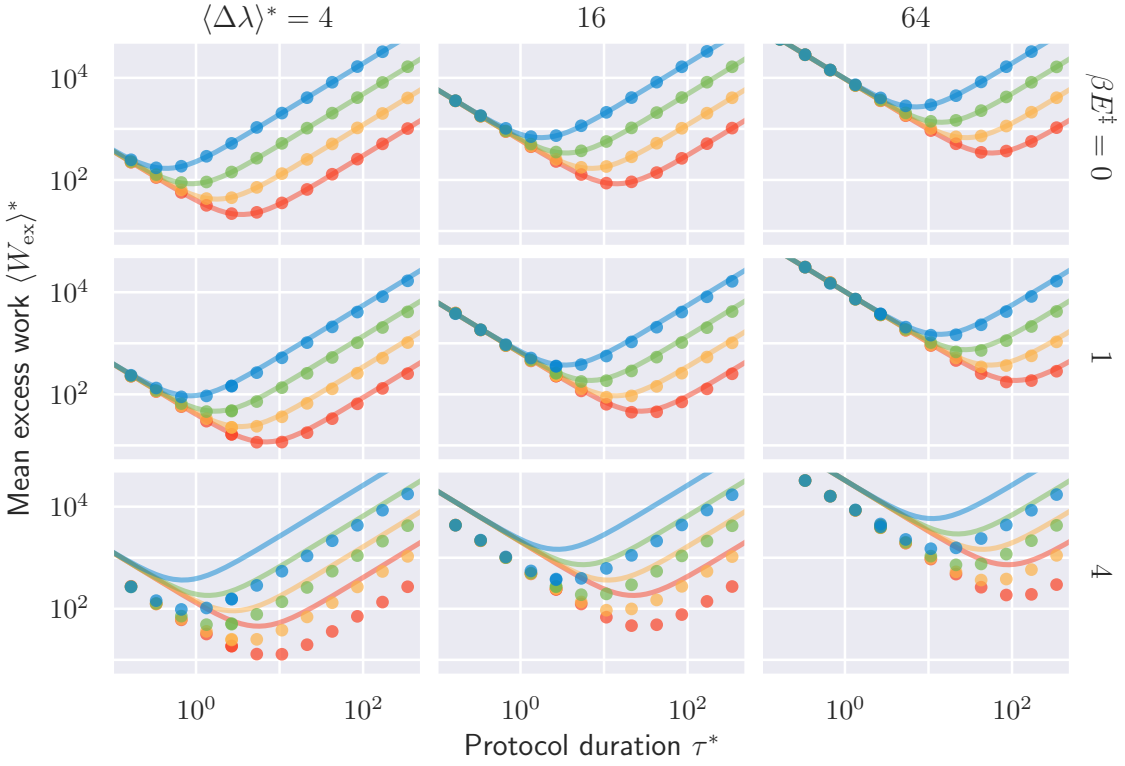

Figure 6.3: **Excess work for periodic-potential ensemble is minimized at finite protocol duration.** Excess work $\langle W_{ex} \rangle^*_\Omega \equiv \beta \langle W_{ex} \rangle_\Omega$ (in units of thermal energy), as a function of protocol duration $\tau^* \equiv \tau/(2\beta D k)^{-1}$ (scaled by the time taken to diffuse the standard deviation $\sqrt{\langle \delta x^2 \rangle_\lambda} \equiv (\beta k)^{-1/2}$ in equilibrium position). Solid lines indicate theoretical predictions (6.49), and circles represent numerical simulation results. Protocol distance $\lambda^* \equiv \lambda/(\beta k)^{-1/2}$ is scaled by $\sqrt{\langle \delta x^2 \rangle_\lambda}$. Nondimensionalized control-parameter velocity variance $\langle \delta \dot{\lambda}^2 \rangle^* \equiv \langle \delta \dot{\lambda}^2 \rangle / (4\beta D^2 k)$ ranges from high (blue) to low (red). Each row shows a different periodic barrier height $\beta E^\ddagger$ from 0 (top) to 4 (bottom). The underlying potential has spatial periodicity $L = 4\sqrt{\langle \delta x^2 \rangle_\lambda}$.

in Sec. 6.4.1, here the fixed-distance boundary conditions imposed on each protocol imply that each protocol travels the same net distance in control-parameter space, but the time taken to complete the protocol is random. The system is initialized in the nonequilibrium steady-state (NESS) for the harmonic trap moving with the (constant) average velocity of the protocol ensemble. Additionally, here we exclusively consider the zero-barrier limit of (6.33) ($\beta E^\ddagger = 0$), where the system is driven by a harmonic potential over a flat underlying potential.

Further, we take the control-parameter dynamics to obey an underdamped Langevin equation (2.32)

$$z = \dot{\lambda} \tag{6.50a}$$

$$m_\lambda \dot{z} = -\partial_\lambda E_\lambda(\lambda, \lambda_0) - \frac{1}{\beta_\lambda D_\lambda} z + \sqrt{\frac{2}{\beta_\lambda^2 D_\lambda}} \xi_\lambda(t) \ , \tag{6.50b}$$

where D_λ is the diffusion coefficient of the control-parameter dynamics, m_λ is the effective mass of the control parameter, and β_λ is the inverse temperature of the heat bath interacting with λ. Details of the numerical implementation of the underdamped dynamics are given in App. A.2.1. The potential energy $E_\lambda(\lambda, \lambda_0(t)) = \frac{1}{2} k_\lambda [\lambda - \lambda_0(t)]^2$ is harmonic with spring constant k_λ confining the control parameter and time-dependent minimum $\lambda_0(t)$. $\lambda_0(t)$ moves with constant velocity, and at steady-state the distribution of control-parameter positions and velocities is stationary in the reference frame comoving with $\lambda_0(t)$. As a result, the average control-parameter velocity is constant, and the average protocol $\langle \Lambda \rangle_\Omega$ is the minimum-work protocol [109]. Furthermore, the steady-state variance of the control-parameter velocity is fixed in the comoving frame by the equipartition theorem [84], $\langle \delta \dot{\lambda}^2 \rangle_{\Omega(t)} = (\beta_\lambda m_\lambda)^{-1}$.

If control-parameter velocity fluctuations (with relaxation time $(\beta_\lambda D_\lambda k_\lambda)^{-1}$) persist over timescales longer than the conjugate-force relaxation time $(\beta D k_{\text{trap}})^{-1}$ of the system, then (6.4) holds for all stochastic protocols in the ensemble.[1] Effectively, this represents the locally smooth limit for control-protocol dynamics, where—over the relaxation timescale of the system—the control parameter is largely unaffected by stochastic fluctuations, but still exhibits large fluctuations over longer timescales.

Figure 6.4 shows the average excess work as a function of the average protocol duration $\langle \tau \rangle_\Omega$, for several (fixed) protocol distances $\Delta\lambda$ and control-parameter diffusion coefficients D_λ. Numerical simulations agree well with the theoretical predictions at short protocol durations, where the excess work is dominated by the contribution from the average protocol (6.32). At long protocol durations, for intermediate-to-large protocol distances and high D_λ (relative to the system diffusion coefficient D), the locally deterministic approximation (6.4) is satisfied and the theoretical predictions agree well with the numerical results. In all cases, the excess work is an increasing function of protocol duration in the long-duration limit. Thus, regardless of the theoretical approximation's accuracy, a finite-time lower bound on the excess work is widely observed, contrary to the case of deterministic protocols.

[1] For a harmonic system, the forces are linear in position $f \propto x$, and thus $\langle \delta f(0) \delta f(t) \rangle \propto \langle \delta x(0) \delta x(t) \rangle$. This implies that the force relaxation time and position relaxation time are equivalent.

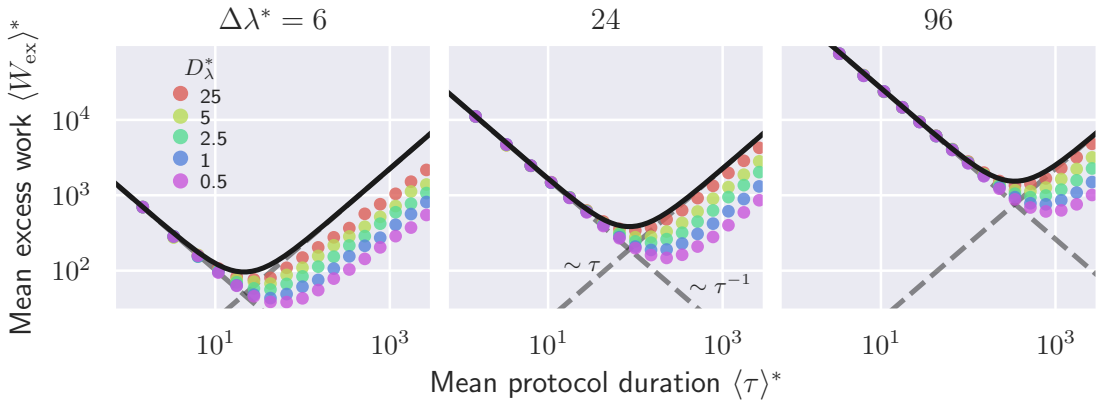

Figure 6.4: **Excess work for stochastic-protocol ensemble matches theoretical approximation in the locally deterministic limit.** Mean excess work $\langle W_{\text{ex}} \rangle^*$ as function of mean protocol duration $\langle \tau \rangle^*$. For underdamped control-parameter dynamics and large protocol distances, hence where control-parameter dynamics are locally deterministic, numerical simulations (circles) agree with the theoretical approximation (6.27) (solid black curve), composed of terms proportional and inversely proportional to protocol duration (dashed gray curves). Control-parameter diffusion coefficient $D_\lambda^* \equiv D_\lambda / D$ (nondimensionalized by the system diffusion coefficient) interpolates between overdamped (purple) and underdamped (red) control-parameter dynamics.

6.5 Discussion

In this chapter, we presented a formalism that generalizes previous theory (Sec. 2.6.2) to now quantify the nonequilibrium costs of driving a system with an ensemble of protocols. We assume only that the linear-response approximation applies for each protocol in the ensemble and that variation about the average protocol is sufficiently small. Furthermore, we have argued that for stochastic control protocols—such as those considered in Sec. 6.4.2—the additional constraint of locally smooth protocols (6.4) is required to ensure the validity of the linear-response approximation. In these limits, protocol variation produces an additional energetic cost that increases with protocol duration.

This theoretical framework gives rise to a lower bound on the excess work (6.29) that generalizes a previous result [180] to arbitrary low-noise protocol ensembles and situations where the relaxation time varies across control-parameter space. Our expression for excess work makes transparent that the lower bound occurs for a finite protocol duration (6.28) and hence finite average protocol velocity (6.30a). This implies an energetically optimal, finite timescale for the process, suggesting the novel possibility that biomolecular processes have energetically preferential timescales over which to operate, stemming from the statistical properties of their driving processes.

The resulting total work is completely specified by the average protocol and the variance of control-parameter velocities (6.27), so may be identical for vastly different control strategies, each with potential advantages for particular tasks. This suggests that an autonomous system could simultaneously reduce the energetic cost of completing a particular thermodynamic process and improve an orthogonal quality metric through the clever choice of the statistical properties of the protocol ensemble (see App. C.3 for more discussion).

We have numerically investigated the consequences of these predictions in two model ensembles. Both the periodic-potential ensemble and stochastic-protocol ensemble show a finite-duration minimum for the excess work across all examined parameter space. Complementary recent experiments [25] have shown that, even far from equilibrium, the linear-response formalism can be effective for reducing excess work in control protocols that unfold and refold a DNA hairpin (see Ch.3-5). Thus, the qualitative trends predicted from our theoretical and numerical investigation may still prove insightful for the operational principles of biomolecular machines, even if such machines' natural operation quantitatively violates linear-response theory. In fact, future experiments (similar to those outlined in Part I) could be designed to more closely match autonomous machines driven by fluctuating forces. Insights from the experiments designed with this framework should provide a deeper understanding of the nonequilibrium energetic efficiency of biomolecular machines and, ultimately, guide the operation of efficient synthetic nanomachines.

This theory is agnostic about the origin of such stochastic control-parameter fluctuations, assigning work to any energy flow during control-parameter changes. Intriguing recent work [181] sheds light on the manner in which nonequilibrium reservoirs can perform work on thermodynamic systems and points toward more biophysically motivated models in which this theory could be applied. Recent research on strongly coupled systems [182] suggests connections with the framework developed here, and the implications for our theory in the broader scope of multiple interacting stochastic systems [183]. However, throughout this chapter, we have made an assumption that the control-parameter dynamics are independent of the stochastic system response, and thus $\boldsymbol{\lambda}$ receives no feedback from the system. This assumption requires further consideration within the context of thermodynamic systems which are detailed balanced. We will revisit this point in Chapter 9.

Chapter 7

Optimal discrete control: minimizing dissipation in discretely driven systems

7.1 Introduction

In Chapter 6, we generalized—in a weak-noise limit—the linear-response excess-work framework [109] to capture additional dissipative costs which arise due to variation in the particular protocol being completed. Our theoretical framework [76] was motivated, in large part, by the stochastic dynamics of molecular machines. Our efforts on stochastic control protocols sought to utilize the theoretical work on optimal control in nonequilibrium systems for understanding the physics of molecular machines. However, there are other ways in which the control protocols discussed in Chapter 6 differ from the way in which biomolecular machines are driven.

As discussed, much of the previous work on the properties of minimum-dissipation control [13, 65, 106, 107, 117, 125, 126, 127, 128, 184] has assumed that the system of interest is subjected to a controlling apparatus that can be manipulated in a continuous manner [109]. While this applies well to the single-molecule experimental paradigm, many interesting microscopic systems, such as biomolecular machines, often drive their mechanical motion via a sequence of chemical reactions. The time scales of chemical reaction and mechanical response can differ by several orders of magnitude; as such, the driving process is well approximated by a series of stochastically timed discrete perturbations to a thermodynamic system, as opposed to a continuous driving process. Furthermore, such discrete perturbations inevitably violate the locally smooth protocol limit required for the results in Chapter 6 to remain valid.

In this chapter, as a first step toward adapting the control-protocol framework to chemically driven systems, and to isolate the effect of discrete control-parameter changes, we develop a theoretical framework for nonequilibrium control using deterministic discrete control-parameter changes. Our central result (7.16) quantifies the nonequilibrium energetic

costs associated with discretely driving a microscopic system. In particular, we assume the system is within the linear-response regime and is subject to sufficiently weak perturbations to warrant a low-order approximation of the energetic costs of discrete steps. Within these limitations, our framework allows for straightforward optimization of discrete driving protocols that accounts for both the effects of the size of discrete steps as well as the local relaxation times, leading to novel characteristics of discrete protocols not observed in continuously driven systems (Fig. 7.5). This work complements and generalizes previous results on the entropy production associated with discrete processes [123]. Furthermore, in the continuous-driving limit, our formalism reduces to previously known results, namely the thermodynamic-length formalism introduced by Crooks [124] and the entropy-differential metric of Burbea and Rao [185], and is related to the generalized friction coefficient of Sivak and Crooks [109].

The chapter proceeds as follows: Sec. 7.2 introduces the relevant theoretical background information; Sec. 7.3 derives an exact expression for the energetic cost of a series of discrete steps applied to an equilibrium system, and then investigates a small-perturbation approximation of this expression to compare to established results; Sec. 7.4 extends this analysis to systems which are out of equilibrium, quantifying within the linear-response regime the excess energetic cost due to incomplete relaxation and in Sec. 7.4.1 discusses how the general result (7.16) relates to the generalized-friction framework introduced in 2.6.2; Sec. 7.5 discusses general implications of optimal, minimum-work protocols in discretely driven systems, under the assumptions laid out in the previous sections; finally, Sec. 7.6 and Sec. 7.7 explore the quantitative implications of this theoretical framework in two model systems: a translating harmonic trap (Sec. 2.7.1) and a periodic potential (Sec. 2.7.2).

7.2 Background

In general, a control protocol $\Lambda : \boldsymbol{\lambda}_0 \to \boldsymbol{\lambda}_N$ is a particular time-dependent transformation applied to the vector of control parameters $\boldsymbol{\lambda} \equiv (\lambda^0, \lambda^1, \cdots, \lambda^n)^T$ to transform it between an initial $\boldsymbol{\lambda}_0$ and final $\boldsymbol{\lambda}_N$ in a prescribed time τ. For a given control protocol, the system responds stochastically. Here, we consider discrete control protocols which consist of a series of instantaneous changes $\Delta\boldsymbol{\lambda}_{i+1,i} \equiv \boldsymbol{\lambda}_{i+1} - \boldsymbol{\lambda}_i$ for consecutive control parameter values $\boldsymbol{\lambda}_i$ and $\boldsymbol{\lambda}_{i+1}$. The system spends a prescribed time Δt_i at each control parameter $\boldsymbol{\lambda}_i$. Thus, each protocol is defined by a set of control-parameter values and the associated times spent at them:

$$\Lambda \equiv \{\lambda_i, \Delta t_i\} \ . \tag{7.1}$$

Previous work has considered energetic flows in discrete-stepping processes [123, 186, 187, 188], but these efforts typically focused on the continuum limit. More recent investigations of driven nonequilibrium systems [106, 109, 117, 126, 127, 128] have focused on control protocols which are continuous functions of time. In contrast to these previous works, and

motivated by the chemically driven paradigm characteristic of microscopic machines, we consider a control protocol as a series of discrete steps of substantial size. However, in Sec. 7.4.1 we discuss how, in the continuum limit, our approach is related to the generalized-friction framework introduced in Sec. 2.6.2.

The average work (divided by $k_{\rm B}T$) associated with a particular discrete perturbation that transforms the control parameter from $\boldsymbol{\lambda}_i$ to $\boldsymbol{\lambda}_{i+1}$ during the protocol Λ is

$$\langle \beta W \rangle_{\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}} = \beta \int \left[E(x, \boldsymbol{\lambda}_{i+1}) - E(x, \boldsymbol{\lambda}_i) \right] p_\Lambda(x, t_{i+1,i}) \, \mathrm{d}x \ , \tag{7.2}$$

where $p_\Lambda(x, t_{i+1,i})$ is the (generally nonequilibrium) distribution over system microstates x at the time $t_{i+1,i}$ that the control parameter changes from $\boldsymbol{\lambda}_i$ to $\boldsymbol{\lambda}_{i+1}$ during the protocol Λ, and angle brackets $\langle \cdots \rangle_{\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}}$ indicate a nonequilibrium average as the control parameter changes from $\boldsymbol{\lambda}_i$ to $\boldsymbol{\lambda}_{i+1}$ during protocol Λ.

We consider protocols which start in equilibrium at initial control-parameter value $\boldsymbol{\lambda}_0$, equivalent to taking the time spent at $\boldsymbol{\lambda}_0$ to infinity, $\Delta t_0 \to \infty$. A particular control protocol begins with the first control-parameter change $\boldsymbol{\lambda}_0 \to \boldsymbol{\lambda}_1$, and finishes when the control parameter $\boldsymbol{\lambda}$ arrives at its terminal value $\boldsymbol{\lambda}_N$. The protocol duration τ is the time taken to complete a given protocol, not counting the time taken to equilibrate at the initial control-parameter value. Thus, for a protocol with $N+1$ control-parameter values $\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N$, the protocol duration is

$$\tau \equiv \sum_{i=1}^{N-1} \Delta t_i \ , \tag{7.3}$$

and the total control-parameter displacement for fixed control-parameter endpoints $\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_N$ is

$$\Delta \boldsymbol{\lambda}_{\rm tot} \equiv \sum_{i=0}^{N-1} \Delta \boldsymbol{\lambda}_{i+1,i} = \boldsymbol{\lambda}_N - \boldsymbol{\lambda}_0 \ . \tag{7.4}$$

For given control-parameter endpoints $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_N$ (hence given protocol displacement $\Delta \boldsymbol{\lambda}_{\rm tot}$) and duration τ, minimizing work involves choosing intermediate control-parameter values $\boldsymbol{\lambda}_i$ and associated dwell times Δt_i.

7.3 Infinite-time work

We first consider the work associated with making a single discrete change to the control parameter of a system which is initially at equilibrium with the control parameter $\boldsymbol{\lambda}_0$ (1.1). The average work required to discretely change the control-parameter vector from $\boldsymbol{\lambda}_0$ to $\boldsymbol{\lambda}_1$ is (7.2), with the equilibrium initial distribution $p_\Lambda(x, t_{1,0}) = \pi(x|\boldsymbol{\lambda}_0)$ (1.1):

$$\langle \beta W \rangle_{\boldsymbol{\lambda}_0 \to \boldsymbol{\lambda}_1} = \beta \int \left[E(x, \boldsymbol{\lambda}_1) - E(x, \boldsymbol{\lambda}_0) \right] \pi(x|\boldsymbol{\lambda}_0) \mathrm{d}x \ . \tag{7.5}$$

From the definition of the equilibrium ensemble at a fixed control parameter (1.1), the energy can be expressed in terms of the equilibrium distribution and free energy, $\beta E(x, \boldsymbol{\lambda}_i) = -\ln \pi(x|\boldsymbol{\lambda}_i) + \beta F(\boldsymbol{\lambda}_i)$. Thus, the average work from (7.5) is given solely in terms of equilibrium distributions,

$$\langle \beta W \rangle_{\boldsymbol{\lambda}_0 \to \boldsymbol{\lambda}_1} = \int \ln \left[\frac{\pi(x|\boldsymbol{\lambda}_0)}{\pi(x|\boldsymbol{\lambda}_1)} \right] \pi(x|\boldsymbol{\lambda}_0) \mathrm{d}x + \beta \Delta F_{1,0} \ , \tag{7.6}$$

where $\beta \Delta F_{1,0} = \beta F(\boldsymbol{\lambda}_1) - \beta F(\boldsymbol{\lambda}_0)$ is the difference between the equilibrium free energy at control parameters $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_1$. The integral in (7.6) is the *relative entropy* (or Kullback-Leibler divergence) $D[p(x)||q(x)] \equiv \int \ln[p(x)/q(x)]\, p(x) \mathrm{d}x$ [94] between the equilibrium distributions before ($p(x) = \pi(x|\boldsymbol{\lambda}_0)$) and after ($q(x) = \pi(x|\boldsymbol{\lambda}_1)$) the control-parameter change,

$$\langle \beta W \rangle_{\boldsymbol{\lambda}_0 \to \boldsymbol{\lambda}_1} = D[\pi(x|\boldsymbol{\lambda}_0) \,\|\, \pi(x|\boldsymbol{\lambda}_1)] + \beta \Delta F_{1,0} \ . \tag{7.7}$$

A protocol consists of N such control-parameter steps $\Delta \boldsymbol{\lambda}_{i+1,i}$, for $i = 0, \ldots, N-1$. If at each control-parameter value $\boldsymbol{\lambda}_i$ the system fully equilibrates, then the average work associated with *any* step $\Delta \boldsymbol{\lambda}_{i+1,i}$ is of the same form (7.7), and the work to complete the entire protocol is the sum of the work associated with each individual step,

$$\langle \beta W \rangle_{\Lambda} = \sum_{i=0}^{N-1} D \left[\pi(x||\boldsymbol{\lambda}_i) \,\|\, \pi(x|\boldsymbol{\lambda}_{i+1}) \right] + \beta \Delta F_{\mathrm{tot}} \ , \tag{7.8}$$

where $\beta \Delta F_{\mathrm{tot}} \equiv \beta F(\boldsymbol{\lambda}_N) - \beta F(\boldsymbol{\lambda}_0) = \sum_{i=0}^{N-1} \beta \Delta F_{i+1,i}$ is the equilibrium free energy change between the initial and final control-parameter values. Thus the average excess work is

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Lambda} \equiv \langle \beta W \rangle_{\Lambda} - \beta \Delta F_{\mathrm{tot}} \tag{7.9a}$$

$$= \sum_{i=0}^{N-1} D[\pi(x|\boldsymbol{\lambda}_i) \,\|\, \pi(x|\boldsymbol{\lambda}_{i+1})] \ . \tag{7.9b}$$

For sufficiently small control-parameter steps $\Delta \boldsymbol{\lambda}_{i+1,i}$, the relative entropy in (7.9b) can be Taylor expanded about its current value $\boldsymbol{\lambda}_i$ to yield

$$D \left[\pi(x|\boldsymbol{\lambda}_i) \,\|\, \pi(x|\boldsymbol{\lambda}_{i+1}) \right] \approx \frac{1}{2} \beta^2 \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i} \Delta \lambda^j_{i+1,i} \Delta \lambda^k_{i+1,i} \ , \tag{7.10}$$

where $\langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i}$ is the equilibrium covariance of conjugate forces $f_j \equiv -\partial_{\lambda^j} E$ at control parameter $\boldsymbol{\lambda}_i$ (see D.1 for details) [109].

Substituting (7.10) in (7.9b) gives the average excess work required to perform the infinite-time discrete protocol Λ,

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Lambda} \approx \frac{1}{2} \beta^2 \sum_{i=0}^{N-1} \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i} \Delta \lambda^j_{i+1,i} \Delta \lambda^k_{i+1,i} \ , \tag{7.11}$$

which—in the continuous-protocol limit—is equivalent to the Burbea-Rao entropy differential metric [185] and the thermodynamic metric derived by Crooks [124].

7.4 Nonequilibrium excess work

To consider the more general situation of finite-time protocols, where at each control-parameter value the system does not fully equilibrate, we appeal to linear-response theory for a step perturbation (see Sec. 2.6.1) [110]. For a system at equilibrium for control parameter $\boldsymbol{\lambda}_{i-1}$, the energy at the next control-parameter value $\boldsymbol{\lambda}_i$ in the protocol Λ can be linearly approximated as

$$E(x,\boldsymbol{\lambda}_i) \approx E(x,\boldsymbol{\lambda}_{i-1}) + \nabla_{\boldsymbol{\lambda}} E(x,\boldsymbol{\lambda})|_{\boldsymbol{\lambda}_{i-1}} \cdot \Delta\boldsymbol{\lambda}_{i,i-1} \tag{7.12a}$$

$$= E(x,\boldsymbol{\lambda}_{i-1}) - \boldsymbol{f}|_{\boldsymbol{\lambda}_{i-1}} \cdot \Delta\boldsymbol{\lambda}_{i,i-1} \tag{7.12b}$$

where $\boldsymbol{f}|_{\boldsymbol{\lambda}_{i-1}}$ is the vector of conjugate forces with elements f_j, evaluated at $\boldsymbol{\lambda}_{i-1}$. When the control parameter instantaneously changes from $\boldsymbol{\lambda}_{i-1}$ to $\boldsymbol{\lambda}_i$, the time-dependent relaxation of f_j towards its equilibrium value at $\boldsymbol{\lambda}_i$ is, under the linear-response approximation,

$$\langle f_j(\Delta t)\rangle_{\boldsymbol{\lambda}_i,\boldsymbol{\lambda}_{i-1}} = \langle f_j\rangle_{\boldsymbol{\lambda}_i} + \beta\langle \delta f_j(0)\delta f_k(\Delta t)\rangle_{\boldsymbol{\lambda}_i}\Delta\lambda^k_{i,i-1} \ , \tag{7.13}$$

where $\langle f_j(\Delta t)\rangle_{\boldsymbol{\lambda}_i,\boldsymbol{\lambda}_{i-1}}$ indicates an average over the instantaneous *nonequilibrium* system distribution after relaxing (under the Hamiltonian $E(x,\boldsymbol{\lambda}_i)$) for a time Δt starting from the equilibrium distribution at $\boldsymbol{\lambda}_{i-1}$. Both the force autocovariance and the average force on the RHS are taken over the *equilibrium* ensemble at fixed control-parameter value $\boldsymbol{\lambda}_i$. Here, we are applying the linear-response framework introduced in Sec. 2.6.1 to the conjugate forces.

If the system relaxes for a time Δt_i at control parameter $\boldsymbol{\lambda}_i$ before the next control-parameter change $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}$, this step requires average work (7.2)

$$\langle \beta W\rangle_{\boldsymbol{\lambda}_i\to\boldsymbol{\lambda}_{i+1}} = \beta\Delta\lambda^j_{i+1,i}\langle f_j(\Delta t_i)\rangle_{\boldsymbol{\lambda}_{i-1},\boldsymbol{\lambda}_i} \tag{7.14a}$$

$$\approx \beta\Delta\lambda^j_{i+1,i}\langle f_j\rangle_{\boldsymbol{\lambda}_i} + \beta^2\Delta\lambda^j_{i+1,i}\langle \delta f_j(0)\delta f_k(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}\Delta\lambda^k_{i,i-1} \tag{7.14b}$$

$$= \langle \beta W^{\infty}\rangle_{\boldsymbol{\lambda}_i\to\boldsymbol{\lambda}_{i+1}} + \langle \beta W^{\mathrm{neq}}_{\mathrm{ex}}\rangle_{\boldsymbol{\lambda}_i\to\boldsymbol{\lambda}_{i+1}} \tag{7.14c}$$

for the linear-force approximation of the infinite-time work $\langle \beta W^{\infty}\rangle_{\boldsymbol{\lambda}_i\to\boldsymbol{\lambda}_{i+1}} = \beta\Delta\lambda^j_{i+1,i}\langle f_j\rangle_{\boldsymbol{\lambda}_i}$ from Sec. 7.3 and linear-response correction $\langle \beta W^{\mathrm{neq}}_{\mathrm{ex}}\rangle$ due to incomplete system relaxation.

Equation (7.14c) is only strictly valid if the system was at equilibrium for $\boldsymbol{\lambda}_{i-1}$ before the step to $\boldsymbol{\lambda}_i$, so more generally the work required for the control-parameter change $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}$ includes contributions from all previous steps. However, (7.14c) approximates the work when the force autocovariance associated with the most recent step is the largest time-dependent contribution to the excess work. This limit is reached when the time spent at each control-parameter value is long compared to the local relaxation time of the conjugate forces. As

discussed in detail in Sec. 7.4.1, this approximation is fundamentally distinct from those made in deriving the continuous-protocol formalism [109].

Within this limit, the total average excess work required to perform a discrete control protocol Λ is

$$\langle \beta W_{\rm ex} \rangle_\Lambda \equiv \sum_{i=0}^{N-1} \langle \beta W_{\rm ex} \rangle_{\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}} \tag{7.15a}$$

$$\approx \langle \beta W_{\rm ex}^{\infty} \rangle_\Lambda + \beta^2 \sum_{i=1}^{N-1} \Delta\lambda_{i+1,i}^j \langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{i,i-1}^k \tag{7.15b}$$

$$\approx \beta^2 \sum_{i=0}^{N-1} \frac{1}{2} \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{i+1,i}^j \Delta\lambda_{i+1,i}^k + \beta^2 \sum_{i=1}^{N-1} \Delta\lambda_{i+1,i}^j \langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{i,i-1}^k \ , \tag{7.15c}$$

where (7.15c) uses the infinite-time excess work for small steps (7.11). The assumption that the system begins in equilibrium at $\boldsymbol{\lambda}_0$ ensures that the $i = 0$ term does not contribute to the total nonequilibrium excess work. By setting $\boldsymbol{\lambda}_{-1} = \boldsymbol{\lambda}_0$, both sums in (7.15) can be taken over the same index range, leading to a more compact form for the excess work:

$$\langle \beta W_{\rm ex} \rangle_\Lambda = \beta^2 \sum_{i=0}^{N-1} \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{i+1,i}^j \Delta\lambda_{i+1,i}^k \left[\frac{1}{2} + \frac{\langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i}}{\langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i}} \frac{\Delta\lambda_{i,i-1}^k}{\Delta\lambda_{i+1,i}^k} \right] \ . \tag{7.16}$$

This captures the combined effects of the control-parameter step sizes $\Delta\boldsymbol{\lambda}_{i+1,i}$ and time allocations Δt_i on the excess work during a discrete control protocol Λ. Here, the time-independent, infinite-time contribution penalizes large control-parameter steps departing from regions with large force covariance, while the time-dependent linear-response correction penalizes steps that are particularly quick (reflected by the force autocorrelation factor) and/or large (reflected by the $\Delta\lambda_{i+1,i}^j \Delta\lambda_{i,i-1}^k$ step-size factor).

Thus, (7.16) generalizes the near-equilibrium expression for the dissipation of a discrete control protocol in [123], because the time dependence captured by the conjugate-force autocovariance in our approach allows for non-exponential relaxation kinetics, and the explicit form permits simultaneous optimization of both the placement of control-parameter values $\boldsymbol{\lambda}_i$ as well as the allocation of times Δt_i. Moreover, Sec. 7.4.1 provides an alternative derivation of the nonequilibrium excess-work contribution (7.14c) using dynamic linear-response theory to show that in the continuous-protocol limit, a linear-response correction to the excess work can recover the generalized-friction formalism from [109], but only if approximations are made which are incompatible with those used to derive (7.16).

7.4.1 Nonequilibrium excess work: linear response for time-dependent protocols

Now, we consider an alternative derivation of the linear-response correction to the excess work for discrete control protocols using linear response for a time-dependent protocol, thereby showing how the general form given in (7.16) relates to the generalized-friction formalism introduced in Sec. 2.6.2 [109].

When subjected to a control protocol Λ, dynamic linear-response theory says that at a time t' after the start of the protocol, the average deviation $\langle \delta f_j(t') \rangle_{\Lambda(t')}$ of the jth component of the conjugate force from its equilibrium value at the current control parameter is [109]

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \int_{-\infty}^{t'} \frac{\mathrm{d}}{\mathrm{d}t''} \left[\langle \delta f_j(0) \delta f_k(t'-t'') \rangle_{\boldsymbol{\lambda}(t')} \right] \lambda^k(t'') \, \mathrm{d}t'' \ . \tag{7.17}$$

Here $\langle \delta f_j(t') \rangle_{\Lambda(t')}$ indicates the average deviation of conjugate forces from their equilibrium value at time t' over system response subject to the protocol Λ, and the integral ranges over the entire previous history of the control protocol Λ.

For a discrete control protocol, the time-dependent history $\lambda^k(t'')$ can be represented by a sum of weighted Heaviside functions

$$\Delta\lambda^k_{i+1,i} \theta(t''-t_{i+1,i}) \equiv \begin{cases} 0 \ , & t'' \leq t_{i+1,i} \\ \Delta\lambda^k_{i+1,i} \ , & t'' > t_{i+1,i} \end{cases} \tag{7.18}$$

reflecting the control-parameter jumps of size $\Delta\boldsymbol{\lambda}_{i+1,i}$ occurring at time $t_{i+1,i}$. Using (7.18), the jth component of the force fluctuation (7.17) becomes

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \int_{-\infty}^{t'} \frac{\mathrm{d}}{\mathrm{d}t''} \left[\langle \delta f_j(0) \delta f_k(t'-t'') \rangle_{\boldsymbol{\lambda}_i} \right] \sum_{n=0}^{i-1} \Delta\lambda^k_{n+1,n} \theta(t''-t_{n+1,n}) \, \mathrm{d}t'' \ , \tag{7.19}$$

where $\boldsymbol{\lambda}_i$ is the current control-parameter value, and the average $\langle \delta f_j(t') \rangle_{\Lambda(t')}$ accounts for the contributions due to all previous steps in the discrete protocol Λ.

Integrating (7.19) by parts, observing that the boundary term is zero if the system begins in thermodynamic equilibrium, and substituting the Dirac delta function for the derivative of the Heaviside function, $\partial_{t''} \theta(t''-t_{n+1,n}) = \delta(t''-t_{n+1,n})$, gives

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \int_{-\infty}^{t'} \langle \delta f_j(0) \delta f_k(t'-t'') \rangle_{\boldsymbol{\lambda}_i} \sum_{n=0}^{i-1} \Delta\lambda_{n+1,n} \delta(t''-t_{n+1,n}) \, \mathrm{d}t'' \ . \tag{7.20}$$

Written out term by term, (7.20) takes the form

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \langle \delta f_j(0) \delta f_k(t'-t_{i,i-1}) \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda^k_{i,i-1} + \langle \delta f_j(0) \delta f_k(t'-t_{i-1,i-2}) \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda^k_{i-1,i-2} + \cdots \ , \tag{7.21}$$

which depends on the times of all previous control-parameter jumps. In order to reach our previous result (7.16), we simply truncate the expansion after the first (leading) term. In this approximation, the average force fluctuation after spending a time $\Delta t_i = t' - t_{i,i-1}$ at control parameter $\boldsymbol{\lambda}_i$ is

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i} \Delta \lambda^k_{i,i-1} \ . \tag{7.22}$$

However, now—instead of truncating the series expansion in (7.21)—we consider the continuous-protocol limit of the entire expansion, where the protocol duration τ is fixed while both the step sizes and the times spent at each control parameter $\Delta t_i \equiv t_{i+1,i} - t_{i,i-1}$ become infinitesimally small, such that $\Delta \lambda^k_{i,i-1} / \Delta t_i \to \dot{\lambda}^k(t)$. Specifically, when spending equal time $\Delta t = \tau/(N-1)$ for each step,[1] (7.21) becomes for $t' = t_{i+1,i}$,

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \lim_{\Delta t, \Delta \lambda_i \to 0} \langle \delta f_j(0) \delta f_k(\Delta t) \rangle_{\boldsymbol{\lambda}_i} \Delta \lambda^k_{i,i-1} + \langle \delta f_j(0) \delta f_k(2\Delta t) \rangle_{\boldsymbol{\lambda}_i} \Delta \lambda^k_{i-1,i-2} + \cdots \tag{7.23}$$

$$= \lim_{\Delta t, \Delta \boldsymbol{\lambda}_i \to 0} \sum_{n=1}^{N} \langle \delta f_j(0) \delta f_k(n\ \Delta t) \rangle_{\boldsymbol{\lambda}_i} \frac{\Delta \lambda^k_{N-n+1,N-n}}{\Delta t} \Delta t \tag{7.24}$$

$$= \int_0^\infty \langle \delta f_j(0) \delta f_k(t) \rangle_{\boldsymbol{\lambda}(t')} \dot{\lambda}^k(t) \, \mathrm{d}t \ . \tag{7.25}$$

Equation (7.25) is the same expression as in [109], where it was subsequently simplified by Taylor expanding the velocity term in the integrand to zeroth order about its current value,

$$\dot{\lambda}^k(t') = \dot{\lambda}^k(t') + \mathcal{O}\left(\ddot{\lambda}^k(t') \right) \tag{7.26}$$

$$\approx \dot{\lambda}^k(t') \tag{7.27}$$

This simplifies (7.25) to

$$\langle \delta f_j(t') \rangle_{\Lambda(t')} \approx \left(\int_0^\infty \langle \delta f_j(0) \delta f_k(t'') \rangle_{\boldsymbol{\lambda}(t')} \mathrm{d}t'' \right) \dot{\lambda}^k(t') \tag{7.28}$$

$$= \zeta_{jk}(\boldsymbol{\lambda}(t')) \dot{\lambda}^k(t') \ , \tag{7.29}$$

where $\zeta_{jk}(\boldsymbol{\lambda}(t'))$ is the generalized friction tensor [109].

In summary, the derivation of the average excess work for an explicitly discrete control protocol (7.16) truncates the dynamic linear-response expression (7.19) after first order, whereas the continuous-protocol result (7.29) includes the influence of all previous perturbations, but approximates the previous control-parameter velocities by the current value.

[1] Here, we need not consider variable time allocations because any continuous control protocol Λ is defined entirely by the ratio $\Delta\lambda/\Delta t \to \mathrm{d}\lambda/\mathrm{d}t$ at each point along the protocol. Put another way, any variation in time allocation can be effectively captured in the continuous limit by variations in displacement.

In essence, this approximation makes similar claims to the discrete truncation in (7.22), in that it assumes that the most recent perturbations are the predominant contributors to the excess work. In light of this, the two derivations can be seen as similar approximations which hold in different circumstances: the former for protocols composed of large discrete steps, and the latter for continuous protocols.

7.5 Minimum-work protocols

We now turn our attention to the strategies for driving a nonequilibrium system through a series of discrete perturbations that minimize the excess work. The nonequilibrium excess work in (7.16) provides a relatively simple expression—within the linear-response approximation—for the energetic cost required to perform a particular discrete control protocol Λ. Although the specific form of a minimum-work protocol depends on the particular system, there are two system-independent special cases which admit simple analytic solutions: the infinite-time limit (Sec. 7.3) where the time-dependent term in (7.16) is negligible, and the case where a single exponential relaxation mode dominates the nonequilibrium excess work.

In general, the excess work can be approximated as

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Lambda} = \sum_{i=0}^{N-1} \langle \beta W_{\mathrm{ex}} \rangle_{\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}} \tag{7.30a}$$

$$\approx \sum_{i=0}^{N-1} \Gamma_{jk}(\boldsymbol{\lambda}_{i-1}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}, \Delta t_i) \Delta\lambda^j_{i+1,i} \Delta\lambda^k_{i+1,i} \tag{7.30b}$$

$$\equiv \sum_{i=0}^{N-1} \mathcal{D}_i^2 \ , \tag{7.30c}$$

for

$$\Gamma_{jk}(\boldsymbol{\lambda}_{i-1}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}, \Delta t_i) \equiv \beta^2 \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i} \left[\frac{1}{2} + \frac{\langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i}}{\langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i}} \frac{\Delta\lambda^k_{i-1,i}}{\Delta\lambda^k_{i+1,i}} \right] \ . \tag{7.31}$$

Here, the k index within the brackets is not summed over. Interpretation is most immediate in the continuous-protocol limit, where each $\mathcal{D}_i$ is the distance along an infinitesimal segment $\mathrm{d}\boldsymbol{\lambda}$ of the control protocol Λ, measured with respect to the metric $\Gamma_{jk}(\boldsymbol{\lambda}_{i-1}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}, \Delta t_i)$; therefore, the sum $\sum_{i=0}^{N} \mathcal{D}_i$ over all steps gives the *thermodynamic length* between the initial and final equilibrium macrostates [123, 124].

For a positive semidefinite force-autocovariance matrix, the total excess work of a particular control protocol can be lower bounded via the Cauchy-Schwarz inequality:

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Lambda} \geq \frac{1}{N} \left(\sum_{i=0}^{N} \mathcal{D}_i \right)^2 \ . \tag{7.32}$$

The lower bound is saturated if and only if the $\mathcal{D}_i$ are identical,

$$\mathcal{D}_i = \mathcal{D} \ . \tag{7.33}$$

Along an optimal protocol (indicated by the superscript $*$), the condition (7.33) implies that $\langle \beta W^*_{\rm ex} \rangle_{\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_{i+1}} = \mathcal{D}^2$, and thus equal excess work is done during each step of the protocol.

For a single control parameter with fixed endpoints λ_0, λ_N and a given set of time allocations Δt_i, the condition (7.33) implies the optimal placement of control parameter values through the proportionality

$$\Delta\lambda^*_{i,i+1} \propto \frac{1}{\sqrt{\Gamma(\lambda^*_{i-1}, \lambda^*_i, \lambda^*_{i+1}, \Delta t_i)}} \ , \tag{7.34}$$

but the implicit dependence of Γ on the step size through $\lambda^*_{i-1}, \lambda^*_i$, and λ^*_{i+1} complicates the practical use of this bound for deriving optimal protocols. However, the proportionality (7.34) can give useful qualitative guidance into the general properties of protocols which saturate the lower bound (7.32). In particular, optimal control-parameter placement tends to avoid regions with large force variance and slowly decaying force autocovariance, subject to the quadratic cost $\Delta\lambda^j_{i+1,i}\Delta\lambda^k_{i+1,i}$ on step sizes. For multiple control parameters, the qualitative insights gained from the lower bound (7.32) and the equality (7.33) can provide a way to derive the optimal time schedule along a particular path in control-parameter space, but unfortunately they do not generally provide a constructive means to identify a path that saturates the bound.

In the infinite-time limit, where $\Gamma_{jk}(\boldsymbol{\lambda}_{i-1}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}, \Delta t_i) \to \Gamma^{(\infty)}_{jk}(\boldsymbol{\lambda}_i) = \frac{1}{2}\beta^2 \langle \delta f_j \delta f_k \rangle_{\boldsymbol{\lambda}_i}$, our predictions reduce to previous calculations by Nulton *et al.* [123] of the optimal placement of discrete steps. In particular, for a single control parameter, the condition (7.33) implies that optimal protocols have the proportionality $\Delta\lambda^*_{i+1,i} \propto 1/\sqrt{\langle \delta f^2 \rangle_{\lambda^*_i}}$. Furthermore, in the continuous-protocol limit, the infinite-time thermodynamic length between the initial and final control parameters (measured with respect to $\Gamma^{(\infty)}_{jk}(\boldsymbol{\lambda}_i)$) converges to that of Crooks [124].

7.5.1 Minimum-work protocols for a single control parameter

For a single control parameter, we perform a constrained optimization of the excess work (7.16),

$$\langle W_{\rm ex} \rangle_\Lambda = \beta \sum_{i=0}^{N-1} \Delta\lambda^2_{i+1,i} \langle \delta f^2 \rangle_{\lambda_i} \left[\frac{1}{2} + \frac{\langle \delta f(0) \delta f(\Delta t_i) \rangle_{\lambda_i}}{\langle \delta f^2 \rangle_{\lambda_i}} \right] \ . \tag{7.35}$$

Here, we use Lagrange multipliers, and follow the method from [123] on a similar problem. First, we define the *Lagrange function*

$$\mathcal{L} = \langle W_{\mathrm{ex}} \rangle_\Lambda - \epsilon_{\mathrm{s}} \left(\sum_{i=1}^{N-1} \Delta t_i - \tau \right) - \epsilon_\tau \left(\sum_{i=0}^{N-1} \Delta\lambda_{i+1,i} - \Delta\lambda_{\mathrm{tot}} \right) \ , \tag{7.36}$$

incorporating fixed protocol endpoints λ_0, λ_N and spatial (7.4) and temporal (7.3) constraints defined in the main text, with respective Lagrange multipliers ϵ_{s} and ϵ_τ.

For the purposes of analytical investigation, we consider naive-space variable-time protocols, where $\Delta\lambda_{i+1,i} = \Delta\lambda_{\mathrm{tot}}/N$ for all steps with fixed endpoints at λ_0, λ_N, and the time allocations Δt_i are variable. We find the optimal allocation of times (indicated by superscript $*$) that minimizes the excess work by extremizing the Lagrange function (7.36) with respect to the allocation of times,

$$0 = \left[\frac{\partial \mathcal{L}}{\partial \Delta t_i} \right]_{\Delta t_i{}^*} \tag{7.37}$$

$$= \beta \Delta\lambda_{i+1,i} \Delta\lambda_{i-1,i} \left[\partial_{\Delta t_i} \langle \delta f_j(0) \delta f_k(\Delta t_i) \rangle_{\boldsymbol{\lambda}_i} \right]_{\Delta t_i{}^*} - \epsilon_\tau \ . \tag{7.38}$$

The brackets $[\cdots]_{\Delta t_i{}^*}$ indicate that the argument is evaluated at $\Delta t_i{}^*$. Without a functional form for $\langle \delta f(0) \delta f(\Delta t_i) \rangle_{\lambda_i}$, this is not analytically tractable.

However, if the autocovariance is dominated by a single exponential relaxation mode (or the time intervals Δt_i are all long enough that the slowest-relaxing mode dominates), then the autocovariance is

$$\langle \delta f(0) \delta f(\Delta t_i) \rangle_{\lambda_i} \approx \langle \delta f^2 \rangle_{\lambda_i} e^{-\Delta t_i / \tau_{\mathrm{R}}(\lambda_i)} \ , \tag{7.39}$$

for the exponential relaxation time $\tau_{\mathrm{R}}(\lambda_i)$. Here, the derivative term in (7.38) is

$$\left[\partial_{t_i} \langle \delta f(0) \delta f(\Delta t_i) \rangle_{\lambda_i} \right]_{\Delta t_i{}^*} = -\frac{\langle \delta f^2 \rangle_{\lambda_i}}{\tau_{\mathrm{R}}(\lambda_i)} e^{-\Delta t_i{}^* / \tau_{\mathrm{R}}(\lambda_i)} \ . \tag{7.40}$$

Substituting this into (7.38) produces

$$\ln(-\epsilon_\tau) = \ln \frac{\beta \Delta\lambda_{i+1,i} \Delta\lambda_{i,i-1} \langle \delta f^2 \rangle_{\lambda_i}}{\tau_{\mathrm{R}}(\lambda_i)} - \frac{\Delta t_i{}^*}{\tau_{\mathrm{R}}(\lambda_i)} \ . \tag{7.41}$$

Summing (7.41) over all steps i in the protocol gives

$$\ln(-\epsilon_\tau) = -\left[\frac{\tau}{\sum_{i=1}^{N-1} \tau_{\mathrm{R}}(\lambda_i)} - \frac{\sum_{i=1}^{N-1} \tau_{\mathrm{R}}(\lambda_i) \ln \beta \mathcal{P}_i}{\sum_{i=1}^{N-1} \tau_{\mathrm{R}}(\lambda_i)} \right] \ , \tag{7.42}$$

where

$$\mathcal{P}_i \equiv \frac{\Delta\lambda_{i+1,i} \Delta\lambda_{i,i-1} \langle \delta f^2 \rangle_{\lambda_i}}{\tau_{\mathrm{R}}(\lambda_i)} \ , \tag{7.43}$$

and we have used (7.3). Equating (7.42) and (7.41), the optimal allocation of time that minimizes the excess work is

$$\Delta t_i^* = \tau_{\mathrm{R}}(\lambda_i)\left[\frac{\tau}{\sum_{n=1}^{N-1}\tau_{\mathrm{R}}(\lambda_n)} - \frac{\sum_{n=1}^{N-1}\tau_{\mathrm{R}}(\lambda_n)\ln(\mathcal{P}_n/\mathcal{P}_i)}{\sum_{n=1}^{N-1}\tau_{\mathrm{R}}(\lambda_n)}\right] . \tag{7.44}$$

This result is equivalent to a similar calculation performed by Nulton *et al.* in [123]. However, our result significantly extends this previous work, as we give a general expression for the magnitude of $\mathcal{P}_i$ in terms of physical quantities (7.43).

In the long-duration limit, where $\tau \gg \sum_{s=1}^{N-1}\tau_{\mathrm{R}}(\lambda_s)$, (7.44) is particularly straightforward, as the second RHS term is negligible, and the optimal allocation of time takes on the simple form

$$\Delta t_i^* \propto \tau_{\mathrm{R}}(\lambda_i) . \tag{7.45}$$

Intuitively, this implies that—in the long-duration limit—minimum-work protocols allocate more time to regions where the integral relaxation time [116] is larger.

The analytical optimization of (7.16) in more complicated scenarios quickly becomes intractable. For instance—even for a single control parameter—using the Lagrange-multiplier method to optimize (7.16) simultaneously with respect to both the control parameter λ_i and the time allocation Δt_i, requires knowledge of partial derivatives of the force autocovariance with respect to both λ_i and Δt_i. Furthermore, these optimizations cannot, in general, be done independently, as the force autocovariance depends on both the time Δt_i as well as the control parameter λ_i.

However, our more general framework (7.16) can be applied to a broader class of problems (as we detail in Sec. 7.6 and Sec. 7.7), in particular to cases with multiple control parameters, non-exponential relaxation kinetics, and optimization of control parameter placements. Thus, for more general scenarios, where analytic optimization methods become cumbersome (and no simple analogs of (7.44) can be found), (7.16) provides a relatively simple expression that can be minimized using numerical methods.

7.6 Harmonic trap

We now focus on a system diffusing in a one-dimensional harmonic trap defined by the potential (Sec. 2.7.1)

$$E_{\mathrm{trap}}(x, \lambda_i) = \frac{1}{2}k_{\mathrm{trap}}(x - \lambda_i)^2 . \tag{7.46}$$

Here k_{trap} is the trap strength and the control parameter is the time-dependent trap minimum λ_i. The work required to perform an N-step discrete control protocol Λ, taking the control parameter from its initial value λ_0 to λ_N, can be calculated exactly.

7.6.1 Infinite-time limit

In the infinite-time limit (Sec. 7.3), the excess work for a single step $\lambda_i \to \lambda_{i+1}$ is the relative entropy (7.7) between the equilibrium distributions (1.1) at λ_i and λ_{i+1}:

$$\begin{aligned} \langle \beta W_{\rm ex} \rangle_{\lambda_i \to \lambda_{i+1}} &= D\left[\pi(x|\lambda_i)||\pi(x|\lambda_{i+1})\right] \\ &= \frac{1}{2}\beta k_{\rm trap} \Delta\lambda_{i+1,i}^2 \ . \end{aligned} \tag{7.47a}$$

Thus the infinite-time work for a discrete protocol of N steps is

$$\langle \beta W_{\rm ex} \rangle_\Lambda = \frac{1}{2}\beta k_{\rm trap} \sum_{i=0}^{N-1} \Delta\lambda_{i+1,i}^2 \ . \tag{7.48}$$

Based on the convexity of this expression, equal step sizes $\Delta\lambda_{i+1,i} = \Delta\lambda_{\rm tot}/N$ minimize the infinite-time work,

$$\langle \beta W_{\rm ex} \rangle_\Lambda \geq \frac{1}{2}\beta k_{\rm trap} \sum_{i=0}^{N-1} \left(\frac{\Delta\lambda_{\rm tot}}{N}\right)^2 = \frac{\beta k_{\rm trap} \Delta\lambda_{\rm tot}^2}{2N} \ , \tag{7.49}$$

which scales with the number of steps as $1/N$ [185]. For this simple system (7.46), the small-step approximation of the relative entropy (7.11) is exact, for arbitrary step sizes.

7.6.2 General solution: finite-time work

Finite-duration control protocols feature both the infinite-time excess work and the time-dependent contribution (Sec. 7.4). For a system initially in equilibrium at the initial control parameter λ_0, the average excess work (in this case equal to the total work since $\Delta F = 0$) for the $(i+1)$th step is

$$\langle \beta W_{\rm ex} \rangle_{\lambda_i \to \lambda_{i+1}} = \beta k_{\rm trap} \Delta\lambda_{i+1,i}^2 \left[\frac{1}{2} + \frac{\xi_{i-1} e^{-\beta D k_{\rm trap} \Delta t_i}}{\Delta\lambda_{i+1,i}}\right] \ , \tag{7.50}$$

where

$$\xi_{i-1} \equiv \textstyle\sum_{n=0}^{i-1} \Delta\lambda_{n+1,n} \exp\left(-\beta D k_{\rm trap} \sum_{r=n+1}^{i-1} \Delta t_r\right) \tag{7.51a}$$

$$= \Delta\lambda_{i,i-1} + \Delta\lambda_{i-1,i-2} e^{-\beta D k_{\rm trap} \Delta t_{i-1}} + \cdots \ . \tag{7.51b}$$

($\xi_{-1} = 0$ as there are no terms in that summation.) Appendix D.2 provides a detailed derivation of (7.50,7.51).

Summing (7.50) over the entire protocol Λ gives

$$\langle \beta W_{\rm ex} \rangle_\Lambda = \beta \sum_{i=0}^{N-1} k_{\rm trap} \Delta\lambda_{i+1,i}^2 \left(\frac{1}{2} + \frac{\xi_{i-1} e^{-\beta D k_{\rm trap} \Delta t_i}}{\Delta\lambda_{i+1,i}}\right) \ . \tag{7.52}$$

For this simple system, the normalized force autocovariance (the force autocorrelation) is $\langle \delta f(0) \delta f(\Delta t_i) \rangle_{\lambda_i} / \langle \delta f^2 \rangle_{\lambda_i} = \exp(-\beta D k_{\text{trap}} \Delta t_i)$, so the approximate excess work within the linear-response regime (7.16) is

$$\langle \beta W_{\text{ex}} \rangle_\Lambda = \beta \sum_{i=0}^{N-1} k_{\text{trap}} \Delta \lambda_{i+1,i}^2 \left(\frac{1}{2} + \frac{\Delta \lambda_{i,i-1}}{\Delta \lambda_{i+1,i}} e^{-\beta D k_{\text{trap}} \Delta t_i} \right) , \tag{7.53}$$

which is equivalent to truncating ξ_{i-1} from (7.51b) after the first term: $\xi_{i-1} \approx \Delta \lambda_{i,i-1}$. From (7.51a), it follows that the linear-response approximation (7.16) holds when

$$\frac{\Delta \lambda_{i+1,i}}{\Delta \lambda_{i,i-1}} \gg e^{-\beta D k_{\text{trap}} \Delta t_i} , \tag{7.54}$$

which is satisfied in the limit of long times $\Delta t_i \gg 1/(\beta D k_{\text{trap}})$ spent at each control-parameter value.

For a protocol Λ of duration τ and consisting of N control parameter steps, each of uniform size $\Delta \lambda_{i+1,i} = \Delta \lambda_{\text{tot}}/N$, with uniform time allocations $\Delta t_i = \tau/(N-1) \equiv \Delta t_{\text{step}}$, the exact excess work is

$$\langle \beta W_{\text{ex}} \rangle_\Lambda = \frac{\beta k_{\text{trap}} \Delta \lambda_{\text{tot}}^2}{N^2} \sum_{i=0}^{N-1} \left(\frac{1}{2} + \frac{\xi_{i-1}}{\Delta \lambda_{i+1,i}} e^{-\beta D k_{\text{trap}} \Delta t_{\text{step}}} \right) , \tag{7.55}$$

the infinite-time work is (7.48), and for large N the linear-response work is

$$\langle \beta W_{\text{ex}} \rangle_\Lambda \approx \frac{\beta k_{\text{trap}} \Delta \lambda_{\text{tot}}^2}{N} \left(\frac{1}{2} + e^{-\beta D k_{\text{trap}} \Delta t_{\text{step}}} \right) . \tag{7.56}$$

In each case, for a fixed duration Δt_{step} allocated to each control-parameter value, the work scales asymptotically ($N \to \infty$) as $1/N$. Figures 7.1a and b show the average excess work for $N = 10$ and $\Delta \lambda_{\text{tot}}/N = 1$, and the difference between the average work and the infinite-time limit as a function of the step duration. For sufficiently large step duration, the exact result (7.55) converges to the linear-response prediction (7.56) and exponentially approaches the infinite-time limit. The three curves have a fixed ordering: the exact solution (7.55) has a series of positive terms added beyond the linear-response expression (7.56), which in turn has an extra positive term added beyond the infinite-time limit (7.49).

For protocol durations sufficiently long that the time spent at each control-parameter value significantly exceeds the relaxation time, the linear-response approximation and the exact result converge. Furthermore, neglecting the ξ_{i-1} term reduces the exact protocol work (7.52) to the infinite-time limit (7.48). Figure 7.1c shows, for the particular step-duration $\Delta t_{\text{step}} = (\beta D k_{\text{trap}})^{-1}$, the $1/N$ scaling (for large N) of the average protocol work.

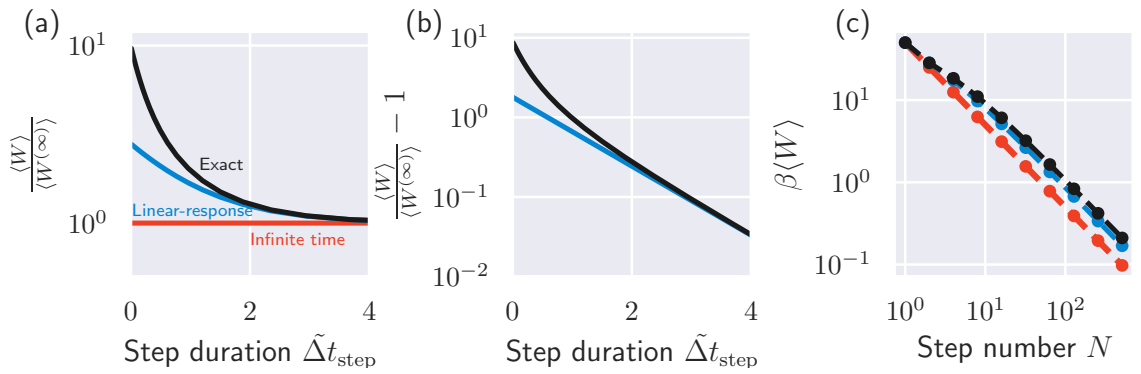

Figure 7.1: **Exact and approximate work for a discretely driven harmonic trap.** Black: exact solution (7.55); blue: linear-response approximation (7.56); red: infinite-time limit (7.49). (a) Excess work, normalized by the infinite-time limit (7.49), as a function of the nondimensionalized step duration $\widetilde{\Delta t}_{\text{step}} \equiv \beta D k_{\text{trap}} \tau/(N-1)$ (scaled by the number of relaxation times spent at each control-parameter value). (b) The difference between the normalized excess work and its infinite-time limit of unity scales exponentially for longer step durations ($\widetilde{\Delta t}_{\text{step}} \gtrapprox 2$), and also converges to the linear-response prediction. (c) Average protocol work $\beta\langle W\rangle$ for a fixed step duration $\widetilde{\Delta t}_{\text{step}} = 1$, as a function of the number N of control parameter steps. The predicted $1/N$ scaling is seen in the exact solution and linear-response approximation at sufficiently large N, and in the infinite-time limit at all N. All plots are for uniform step spacing, $\Delta\lambda_{i+1,i} = \Delta\lambda_{\text{tot}}/N$ for each step i.

7.7 Periodic potential

Now we consider a single diffusing particle in a one-dimensional energy landscape $E_{\text{mol}}(x, \lambda_i)$ consisting of two components: a control-parameter-independent periodic potential, and a control-parameter-dependent harmonic trapping potential (Sec. 2.7.2):

$$E_{\text{periodic}}(x, \lambda_i) = \frac{1}{2} k_{\text{trap}} (x - \lambda_i)^2 - \frac{1}{2} E^{\ddagger} \cos \frac{2\pi x}{\ell} \ , \tag{7.57}$$

where k_{trap} is the harmonic trap strength, $E^{\ddagger}$ is the energetic barrier height between adjacent minima on the periodic potential, and ℓ is the period (Fig. 7.2). This potential represents a system with a sequence of metastable states, such as those often found in models of molecular machines [132], and for $E^{\ddagger} = 0$ reduces to the harmonic potential considered in Sec. 7.6.

Figure 7.2b shows numerical estimation of the autocovariance $\langle \delta f(0) \delta f(\Delta t)\rangle_{\lambda}$ from equilibrium simulations at several fixed control-parameter values evenly spaced over a single period of the underlying potential (7.57). Specifically, we use an overdamped Langevin equation (2.31) for a particle in the potential (7.57) with an underlying period ℓ to model the dynamics of the system. We numerically integrate the equation of motion using the overdamped integrator outlined in App. A.2.2. Given this equation of motion, we calculate the force autocovariance $\langle \delta f(0) \delta f(\Delta t)\rangle_{\lambda}$ over a discrete mesh of control-parameter values

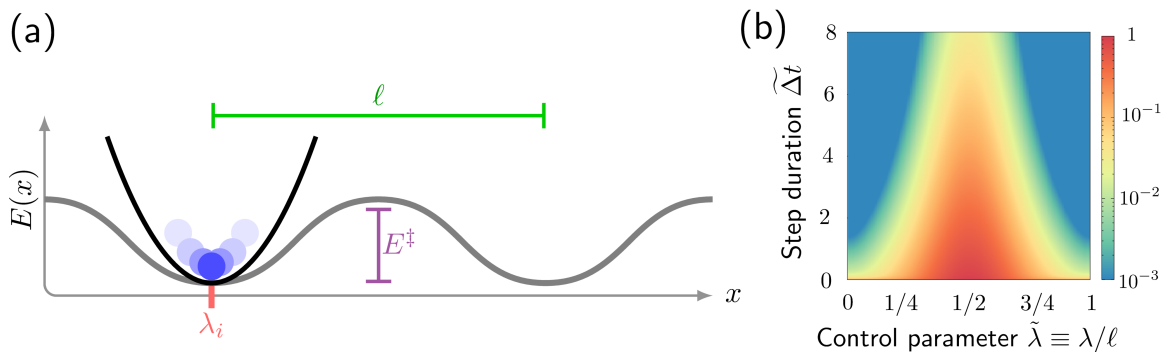

Figure 7.2: **(a) Schematic of the periodic potential.** The control parameter λ_i (red) takes discrete steps $\Delta\lambda_{i+1,i}$ to drive the system (fluctuating blue ball) over a series of energy barriers (of height $E^\ddagger$) separating adjacent metastable potential wells. The underlying potential has period ℓ (green). **(b) Force autocovariance sampled across a single period of the underlying potential.** Heat map for scaled force autocovariance $\langle \delta f(0) \delta f(\Delta t) \rangle_\lambda / \max_\lambda(\langle \delta f^2 \rangle_\lambda)$ as a function of control parameter $\tilde{\lambda} \equiv \lambda/\ell$. $\widetilde{\Delta t} \equiv \Delta t / \overline{\tau_{\text{relax}}}$ is the nondimensionalized step duration, and $\overline{\tau_{\text{relax}}}$ is the period-averaged integral relaxation time.

$\lambda \in [0, \ell)$ and lag times $\Delta t \in (0, \infty)$ by simulating equilibrium dynamics. For the purposes of numerical optimization, we approximate the autocovariance function between mesh points using a bivariate cubic spline interpolation on the empirical $(\lambda, \Delta t)$ mesh. This interpolation scheme yields an approximation of the force autocovariance surface which is smooth, and has continuous first partial derivatives, thus satisfying the necessary smoothness constraints for the numerical optimizer [189]. Using the force autocovariance as input to the linear-response approximation (7.16), we minimize the average excess work during a discrete control protocol Λ with a fixed number N of steps.[2]

We consider three different protocol optimization schemes in order to isolate the effects of the optimal allocation of times Δt_i to a fixed 'naive' sequence of control-parameter values (a 'time-optimized' protocol), the optimal placement of control-parameter values for a fixed 'naive' set of time allocations ('space-optimized'), and the simultaneous optimization of time allocations and control-parameter placements ('fully optimized'). In all cases, protocols are constrained by having fixed protocol duration τ (7.3), protocol endpoints λ_0, λ_N, and number of steps N. In order to minimize the effect of the boundary conditions, we consider control protocols which traverse several periodic repetitions of the underlying potential.

For such a discrete protocol, Fig. 7.3 shows the time allocations Δt_i and control-parameter step-sizes $\Delta\lambda_{i+1,i}$, relative to their naive values $\Delta t_{\text{naive}} \equiv \tau/(N-1)$ and $\Delta\lambda_{\text{naive}} \equiv \Delta\lambda_{\text{tot}}/N$, as a function of the control-parameter value λ over a single period, for a protocol

[2] In particular, we use a quadratic optimization algorithm, implemented using the `SLSQP` (Sequential Least SQuares Programming) option in the `scipy.optimize.minimize` python package.

with $N_\mathrm{p} = 6$ steps per potential period and protocol duration $\tau = 4(N-1)\overline{\tau_\mathrm{relax}}$, where $\overline{\tau_\mathrm{relax}}$ is the mean integral relaxation time over a single period of the underlying potential.

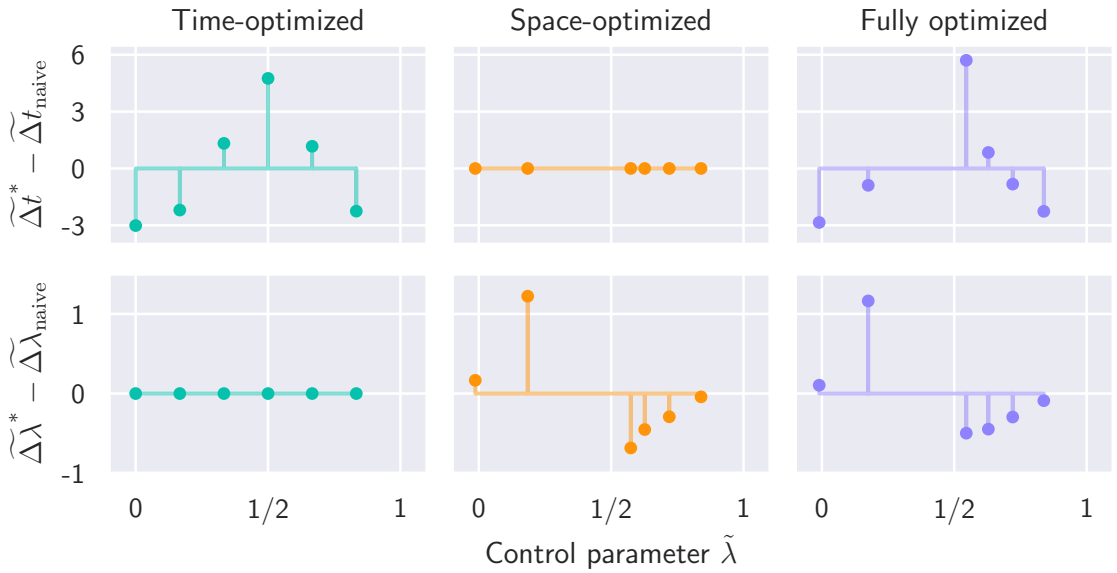

Figure 7.3: **Protocols designed to minimize work allocate time and/or steps significantly differently from naive protocols.** Relative (nondimensionalized) time allocations $\widetilde{\Delta t}^* - \widetilde{\Delta t}_\mathrm{naive}$ (top row) and relative step sizes $\widetilde{\Delta\lambda}^* - \widetilde{\Delta\lambda}_\mathrm{naive}$ (bottom row), for time-optimized (left column), space-optimized (middle column), and fully optimized discrete protocols (right column), with $N_\mathrm{p} = 6$ steps per periodic repetition of the underlying landscape. $\Delta t_\mathrm{naive} \equiv \tau/(N-1)$ is the naive time allocation, and $\Delta\lambda_\mathrm{naive} \equiv \Delta\lambda_\mathrm{tot}/N$ is the naive control-parameter step size. Time allocations are nondimensionalized as $\widetilde{\Delta t} \equiv \Delta t/(N\overline{\tau_\mathrm{relax}})$, where $\overline{\tau_\mathrm{relax}}$ is the mean integral relaxation time over a single period of the underlying potential. Control-parameter step sizes are nondimensionalized as $\widetilde{\Delta\lambda} \equiv \Delta\lambda/(N_\mathrm{p}\ell)$, where N_p is the number of steps in a periodic repetition of the underlying potential, and ℓ is the period of the underlying potential. Protocol has a total duration $\tau = 4(N-1)\overline{\tau_\mathrm{relax}}$.

In each case, the behavior predicted by our theoretical analysis of simplified systems in Sec. 7.5 is borne out. In particular, time-optimized protocols allocate a larger fraction of the protocol duration to regions where the force is slowly relaxing (7.45), while space-optimized protocols take step sizes which are largest in regions where the force variance is small and rapidly relaxing. The fully optimized protocols show both behaviors.

Figure 7.4 shows the theoretically expected excess work for these minimum-work protocols, specifically the predicted excess-work ratio $\langle W_\mathrm{ex}^*\rangle_\Lambda / \langle W_\mathrm{ex}^\mathrm{naive}\rangle_\Lambda$ for the three distinct protocol classes: time-optimized, space-optimized, and fully optimized. For short durations ($\widetilde{\Delta t} \equiv \tau/(N\overline{\tau_\mathrm{relax}}) < 1$), time optimization yields no gain over naive protocols, while spatial optimization and full optimization are indistinguishable. For intermediate durations ($\widetilde{\Delta t} \approx 2$), time optimization has maximum effect, and full optimization significantly improves upon spatial optimization. For longer durations ($\widetilde{\Delta t} \gtrapprox 8$), time optimization again

gives no benefit over the naive protocol, as the time-dependent term in (7.16) becomes negligible.

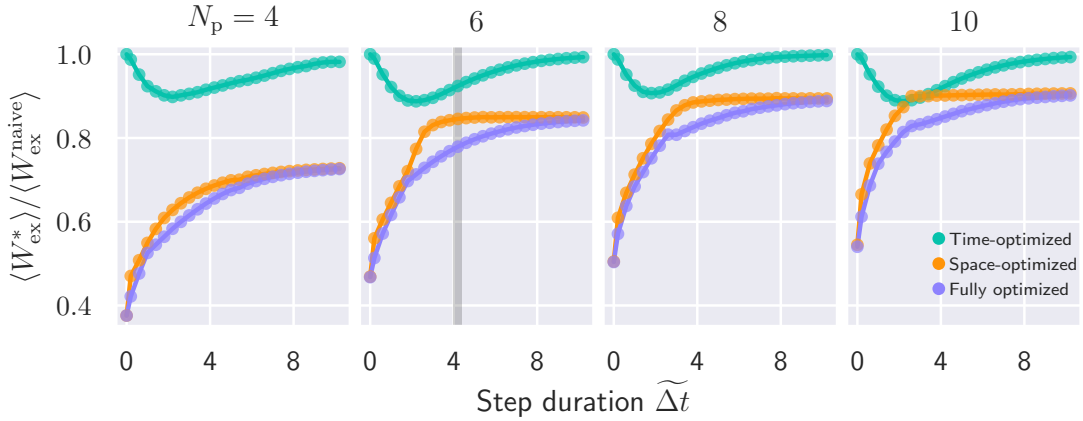

Figure 7.4: **Optimized discrete control protocols significantly reduce the predicted excess work**. Excess-work ratio for discrete control protocols traversing several periodic images, as a function of nondimensionalized step duration $\widetilde{\Delta t} \equiv \tau/(N\overline{\tau_{\text{relax}}})$. Purple: fully optimized protocols (7.16); turquoise: time-optimized; orange: space-optimized. Number N_p of steps per potential period varies from left to right sub-plots. The grey bar on the $N_p = 6$ subplot indicates the protocols shown in Fig. 7.3.

Figure 7.5 shows that as the number N_p of steps per periodic image increases, the time allocation for fully optimized discrete control protocols converges to that of the optimal continuous protocol derived from the generalized friction coefficient [109].

However, there is a significant difference between the discrete (at low step numbers N_p) and continuous control protocols [106, 109]. In particular, relative to an optimal continuous protocol, fully optimized discrete protocols allocate a smaller fraction of their duration at $\lambda = \ell/2$, near the energy barrier (and in fact at the lowest step numbers completely avoid this region). This ability of low step-number protocols to entirely avoid regions of control-parameter space with high force variance and slow relaxation (generally speaking, near energy barriers) represents a qualitatively distinct optimization strategy that is simply not available to continuous protocols.

Furthermore, the continuous protocols allocate time symmetrically about the energy barrier because the generalized friction maintains the same symmetries as the underlying energetic landscape (7.57) [109]. As a result, a continuous optimal protocol traverses the same path in both the forward $\lambda_0 \to \lambda_N$ and reverse $\lambda_N \to \lambda_0$ directions. Discrete protocols break this symmetry because of the infinite-time contribution (Sec. 7.3), which is given by a sum of relative entropies $\langle \beta W_{ex}^{\infty} \rangle = \sum_{i=0}^{N-1} D[\pi(x|\lambda_i)||\pi(x|\lambda_{i+1})]$. In general, $D[p(x)||q(x)] \neq D[q(x)||p(x)]$ and thus the infinite-time excess work $\langle \beta W_{ex}^{\infty} \rangle_{\Lambda}$ is asymmetric in the protocol direction. For small steps (7.11), this asymmetry persists; in (7.16) the excess work during the control-parameter step $\Delta\lambda_{i+1,i}$ is a function of the force variance

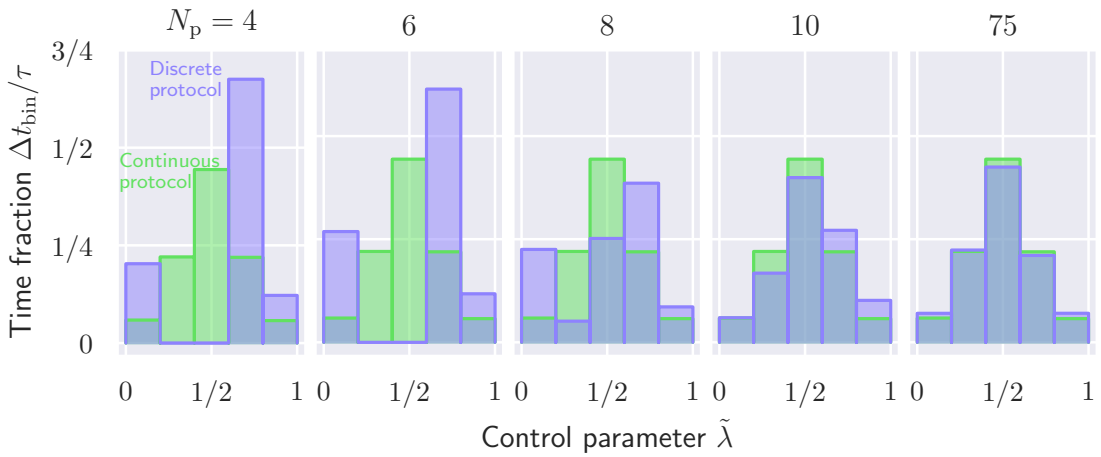

Figure 7.5: **In the many-step/continuous-protocol limit, fully optimized discrete protocols allocate time the same as their continuous-protocol analogs.** The fraction of the total protocol duration spent in each region of control-parameter space for a fully optimized discrete control protocol (green), compared to the fraction during an optimized continuous protocol generated using the generalized-friction framework [109]. As the number N_p of control-parameter values per potential period increases from $N_p = 4$ to $N_p = 75$, the discrete-protocol histogram converges to the continuous-protocol histogram.

at the *current* control parameter $\langle \delta f^2 \rangle_{\lambda_i}$ (and independent of the force variance $\langle \delta f^2 \rangle_{\lambda_{i+1}}$ at the destination control-parameter value λ_{i+1}). This produces a directional asymmetry as the excess work for the control-parameter step $\Delta\lambda_{i,i+1}$ is generally different than the excess work for step $\Delta\lambda_{i+1,i}$. However, as the number of control-parameter steps increases and the distance between those steps becomes sufficiently small, the difference between the force variance at consecutive control-parameter values becomes negligibly small, and the asymmetry between forward and reverse protocols vanishes (Fig. 7.5).

7.8 Discussion

In this chapter, we derived the work required to drive a microscopic system out of equilibrium via a discrete control protocol. Such a control protocol transforms the energy landscape through a series of discrete intermediate states, capturing the discrete nature of the chemical reaction sequences that drive many biological molecular motors. Our central result is the linear-response expression for excess work (7.16), which quantifies the near-equilibrium work of a particular control protocol, solely as a function of the equilibrium system properties.

We deduced a general expression for the work required to make a discrete change in the control-parameter vector of a system in equilibrium (7.9b) and used this to exactly quantify the work required to perform a discrete protocol in the infinite-time limit (Sec. 7.3). When each step is sufficiently small and hence each perturbation is sufficiently weak, our derivation reduces to previously known results [124, 185]. Our primary contribution is to

generalize these analyses beyond the infinite-time limit, where we use a linear-response approximation to derive the leading-order time-dependent contribution to the excess work (Sec. 7.4). Theoretically, our results go significantly beyond previous efforts [123] to quantify energy flows in discretely driven nonequilibrium systems, by incorporating the effects of relaxation kinetics that are non-exponential and that vary across control-parameter space, and by simultaneously optimizing both the placement of control-parameter values as well as the allocation of times.

We investigated the correspondence between our linear-response approximation and an exact solution for a harmonically trapped Brownian particle driven by a series of discrete steps of equal size. We also studied the optimal allocation of time and placement of control-parameter values that minimize the work for protocols traversing many repetitions of a periodic energy landscape (7.57). Here, we find that fully optimized discrete control protocols have qualitatively distinct features when compared to their continuous-protocol analogs. In particular, discrete protocols do not obey the same directional symmetry that continuous protocols do, and in the context of the periodic potential (Sec. 7.7), discrete protocols allocate a smaller fraction of their total duration near the energy barrier. More generally, minimum-work protocols allocate more time to regions where the force has a smaller variance and is more slowly decaying – as discussed in Sec. 7.5. Finally, we quantified the reduction in excess work—relative to a naive protocol—achieved by these minimum-work discrete control protocols. In particular, the theoretical excess-work reduction (relative to a naive protocol) of a fully optimized protocol exceeds 50% for small step numbers and short protocol durations. Significant reduction persists even for intermediate durations ($\tilde{\tau} \approx 2$) when fully optimized, highlighting the benefits of both optimized placement of control-parameter values and the allocation of time among them.

This paradigm of discretely driven nonequilibrium systems is motivated, in part, due to its resemblance to the chemical driving in many molecular machines. However, we have yet to incorporate the stochastic nature of the chemical processes into the discrete driving protocols. In the following chapter (Ch. 8) we briefly discuss the extension of the discrete excess work (7.16) to stochastic protocol ensembles—in a similar approach to Ch. 6—and explore the consequences analytically in the same harmonic model systems considered in Sec. 7.6. This exploration allows us to elucidate the connections with previous work on the dissipation in autonomous systems [180, 190], and understand more clearly the relationship between the dissipation bound derived in Ch. 6 and Machta's dissipation bound in [180].

Chapter 8

On dissipation bounds: discrete stochastic control of nonequilibrium systems

8.1 Introduction

In the previous two chapters (Ch. 6 and 7) we discussed two different generalizations of the linear-response excess-work approximation (introduced in Sec. 2.6.1) for driven, nonequilibrium systems. In particular, we investigated separately the effects of stochastic control-protocol ensembles (Ch.6), and discrete control protocols (Ch. 7). In this chapter, we briefly discuss how to incorporate both of these generalizations into a common framework. As a result, we make a direct connection between Ch. 6 and recent results regarding lower bounds on the dissipation in stochastically driven systems [180, 190], and discuss the implications of these results in molecular machines.

In this chapter we generalize the theoretical framework for discretely driven nonequilibrium systems (Ch. 7) by incorporating—in a similar manner to Ch. 6—the stochasticity of control-parameter dynamics characteristic of a chemical reaction network [88]. Here, we find a novel characterization of the dissipative costs associated with discrete stochastic control which—similar to Ch. 7—decomposes the dissipation into contributions that arise due to the discrete perturbations and those due to the nonequilibrium nature of the system. Additionally, we find a lower dissipation bound that arises from a trade-off between these competing contributions, and provides further insight into the connections between the lower bound in [180, 190] and the bound derived in Ch. 6. We accomplish this by including a dissipative *cost of control* due to the symmetry-breaking of the control-parameter dynamics. Furthermore, we illustrate our results with a simple model system that admits a tractable form for all contributions to the dissipation, and discuss the limits in which the existing dissipation bounds are valid.

8.2 Discrete stochastic protocols

Throughout this chapter we consider control-parameter dynamics that are both discrete and stochastic, such as those found, for instance, in chemical reaction networks [88]. Specifically, we assume control-parameter dynamics are Markovian and thus governed by a master equation 2.2.1, so that the rate of change of the probability distribution of control parameters as a function of time is [72]

$$\mathrm{d}_t p_i = \sum_j V_{ij} p_i \ , \tag{8.1}$$

where p_i is the probability that, at time t, the control parameter is $\boldsymbol{\lambda}_i$, and V_{ij} is the i,jth element of the transition rate matrix, quantifying the rate of transition $\boldsymbol{\lambda}_j \to \boldsymbol{\lambda}_i$. Here, we take each distinct control-parameter vector $\boldsymbol{\lambda}$ to be associated with a unique index i.

Given the nature of our analysis, we are concerned with the response of the system to a particular discrete control protocol—or sequence of control-parameter values and associated lag times—and thus are interested in the characteristics of single trajectories. Trajectories which are consistent with (8.1) (such as those shown in Fig. 2.1) can be generated through a kinetic Monte Carlo scheme [133]. Consistent with our previous assumptions in Chapters 6 and 7, we take the control-parameter dynamics to be independent of the system response, so that the probability of any particular control-parameter trajectory can be captured completely by the transition rates V_{ij}. [1]

Similarly to Ch. 6, we are interested in control-protocol ensembles $\Omega \equiv \{\Lambda\}$ and the associated probabilities $p[\Lambda|\Omega]$ of each individual protocol Λ given a specification of the control-parameter dynamics (through the transition rates V_{ij}) and their boundary conditions [76]. In this chapter, we consider fixed-distance protocol ensembles, where the final location $\boldsymbol{\lambda}_{N_\Omega}$ of each control protocol is fixed, and each protocol has a random duration τ_Λ.

For any particular protocol Λ, the excess work can be approximated by (7.16). Thus, the protocol-ensemble average excess work is

$$\langle W_{\mathrm{ex}} \rangle_\Omega = \int \langle \beta W_{\mathrm{ex}} \rangle_\Lambda \, p[\Lambda|\Omega] \, \mathcal{D}[\Lambda] \ . \tag{8.2}$$

For a particular transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, the stochastic nature of control-parameter dynamics manifests through the the distribution of dwell times in state $\boldsymbol{\lambda}_i$. Thus, at the level of a given transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, to evaluate the protocol-ensemble average, we simply integrate (7.16) for a single step over the distribution of dwell times $p_{\mathrm{dwell}}(\Delta t_i)$:

$$\langle \beta W_{\mathrm{ex}} \rangle_{ji|\ell} = \int \langle \beta W_{\mathrm{ex}}(\Delta t_i) \rangle_{ji|\ell} \, p_{ji}^{\mathrm{dwell}}(\Delta t_i | \boldsymbol{\lambda}_i) \, \mathrm{d}\Delta t_i \ . \tag{8.3}$$

[1] This assumption is also made in Refs. [180, 190], and we will explore its implications in Part III. Specifically, it violates the local detailed-balance constraint on the microscopic rates, and thus represents what we will call a *thermodynamically incomplete* system (Ch. 9).

Here, the angle brackets $\langle\cdots\rangle_{ji|\ell}$ indicate an average over many repetitions of the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, given that the previous control-parameter jump was $\boldsymbol{\lambda}_\ell \to \boldsymbol{\lambda}_i$. Unlike in Ch. 7, where the index i on the control parameter was sequential—increasing monotonically and indicating the relative position of a particular control parameter within the protocol—here the indices i on the control-parameter vectors indicate a specific value, and thus we require the explicit conditioning on the previous state $\boldsymbol{\lambda}_\ell$.

In many cases, the distribution of dwell times is exponential, $p_{ji}^{\mathrm{dwell}}(\Delta t_i) = V_{ji}\exp(-V_{ji}\Delta t_i)$. Additionally, because the infinite-time work $\langle\beta W_{\mathrm{ex}}^{\infty}\rangle_{ji}$ is independent of the step duration Δt_i, only the nonequilibrium excess work $\langle\beta W_{\mathrm{ex}}^{\mathrm{neq}}\rangle_{ji|\ell}$ is affected by the randomness of dwell times. Thus the protocol-ensemble average nonequilibrium excess work is

$$\langle\beta W_{\mathrm{ex}}^{\mathrm{neq}}\rangle_{ji|\ell} \approx \beta^2 V_{ji}\Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}\int\langle\delta f_k(0)\delta f(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}\, e^{-V_{ji}\Delta t_i}\,\mathrm{d}\Delta t_i \ . \tag{8.4}$$

In a moderate timescale-separated limit, where the control-parameter dynamics are significantly slower than the conjugate-force relaxation—which is consistent with the linear-response approximations taken to arrive at (7.16)—the exponential term in (8.4) can be expanded in a Taylor series, and approximated to zeroth order as $\exp(-V_{ji}\Delta t_i) \approx 1$. Thus, the approximate nonequilibrium excess work is

$$\langle\beta W_{\mathrm{ex}}^{\mathrm{neq}}\rangle_{ji|\ell}^{\Omega} \approx \beta^2 V_{ji}\Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}\int_0^\infty\langle\delta f_k(0)\delta f_{k'}(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}\,\mathrm{d}\Delta t_i \tag{8.5a}$$

$$= \beta V_{ji}\Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}\zeta_{kk'}(\boldsymbol{\lambda}_i) \ , \tag{8.5b}$$

where $\zeta_{kk'}(\boldsymbol{\lambda}_i) \equiv \beta\int_0^\infty\langle\delta f_k(0)\delta f_{k'}(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}\,\mathrm{d}\Delta t_i$ is the generalized friction tensor (2.65) originally used to quantify the dissipation in systems driven by deterministic, continuous control protocols.

We can also generalize the calculation of the protocol-ensemble average work for a particular step in (8.5b) to non-exponential dwell-time distributions. This is the case, for instance, in so-called semi-Markov processes [191], such as when multiple stochastic events must occur for a particular control-parameter transition to occur [192]. Here, dwell times are Gamma distributed [193],

$$p_{ij}^{\mathrm{dwell}}(\Delta t) = \frac{(V_{ji})^{\alpha}}{\Gamma(\alpha)}(\Delta t)^{\alpha-1}e^{-V_{ji}\Delta t}\mathrm{d}\Delta t \ , \tag{8.6}$$

where $\alpha > 0$ is a shape parameter and $\Gamma(\alpha)$ is the Gamma function evaluated at α. This two-parameter $(\alpha,\, V_{ji})$ distribution is a generalization of the exponential distribution; for $\alpha = 1$ the Gamma distribution is equivalent to the exponential distribution, for $\alpha < 1$ the dwell times are, on average, smaller than they would be if they were exponentially distributed, and for $\alpha > 1$ the dwell times are, on average, larger. Appendix E.1 derives the nonequilibrium excess work for Gamma-distributed dwell times, which takes a similar form

to (8.5b):

$$\langle \beta W_{\rm ex}^{\rm neq} \rangle_{ji|\ell} = \beta (V_{ji})^{\alpha} \Delta\lambda_{ji}^{k} \Delta\lambda_{i\ell}^{k'} \zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}_i) \tag{8.7}$$

where $\zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}_i) \equiv \frac{\beta}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \langle \delta f_k(0) \delta f_{k'}(t) \rangle_{\boldsymbol{\lambda}} \, \mathrm{d}t$ is a generalization of the friction tensor $\zeta_{kk'}(\boldsymbol{\lambda})$ to Gamma-distributed dwell times. For $\alpha = 1$, (8.7) reduces to the generalized friction tensor in (2.65). In what follows, we will use the friction from (8.7), which captures the typical case of exponential relaxation at $\alpha = 1$.

Substituting (8.5b) into (7.16) for a single control-parameter step gives the protocol-ensemble averaged excess work for a system with stochastic jump dynamics:

$$\langle \beta W_{\rm ex} \rangle_{ji|\ell} \approx \frac{\beta^2}{2} \langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^{k} \Delta\lambda_{ji}^{k'} + \beta V_{ji} \zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}_i) \Delta\lambda_{ji}^{k} \Delta\lambda_{i\ell}^{k} \, . \tag{8.8}$$

Furthermore, because the control-parameter jump dynamics are Markovian, we can write the average excess work associated with the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ (as opposed to the conditional average excess work (8.8) conditioned on the previous control-parameter value) by summing over all possible previous control-parameter values $\boldsymbol{\lambda}_\ell$, with each contribution weighted by $V_{i\ell}/\tilde{V}_i^{\rm tot}$, where $\tilde{V}_i^{\rm tot} \equiv \sum_s V_{is}$ is the total entry rate into state $\boldsymbol{\lambda}_i$ from connected states. Specifically, the sum over index s includes all states for which $V_{is} \neq 0$. Thus, the unconditioned protocol-ensemble averaged excess work for the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ is

$$\langle \beta W_{\rm ex} \rangle_{ji} \approx \frac{\beta^2}{2} \langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^{k} \Delta\lambda_{i\ell}^{k'} + \frac{\beta V_{ji}}{\tilde{V}_i^{\rm tot}} \zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}_i) \Delta\lambda_{ji}^{k} \sum_s V_{is} \Delta\lambda_{is}^{k'} \, . \tag{8.9}$$

Thus—given an ensemble of trajectories of the control-parameter vector—if the generalized friction $\zeta_{kk'}(\boldsymbol{\lambda})$ and force variance $\langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i}$ are known at each of the control-parameter vectors, we can use (8.9) to predict the average dissipation due to the response of the system.

We further condense (8.9) as

$$\langle \beta W_{\rm ex} \rangle_{ji} \approx \beta \Delta\lambda_{ji}^{k} \Delta\lambda_{ji}^{k'} \left[\frac{\beta}{2} \langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i} + \frac{V_{ji}}{\tilde{V}_i^{\rm tot}} \zeta_{kk'}(\boldsymbol{\lambda}_i) \sum_s V_{is} \frac{\Delta\lambda_{is}^{k'}}{\Delta\lambda_{ji}^{k}} \right] \tag{8.10}$$

which makes more transparent that—in the timescale-separated limit where $V_{ji} \to 0$—the infinite-time contribution dominates, and the protocol-ensemble averaged excess-work contribution simplifies to

$$\langle \beta W_{\rm ex} \rangle_{ji} \approx \frac{\beta^2}{2} \langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^{k} \Delta\lambda_{ji}^{k'} \, . \tag{8.11}$$

In this limit, one can derive lower bounds on the excess work associated with a particular protocol Λ through optimally placing the intermediate states, as in Sec. 7.5 [77, 123, 187].

8.3 A cost for control

So far, our formalism for quantifying dissipation associated with discrete stochastic driving protocols has remained agnostic on the origins of stochasticity in the protocols themselves. Furthermore, we have neglected any essential costs associated with generating the control-parameter dynamics. Motivated by complementary works on the dissipative costs of stochastic control protocols [180, 190], we now introduce an effective *cost of control* based on symmetry-breaking in the dynamics of the control protocols.

In general, the dissipative cost associated with generating dynamics that break time-reversal symmetry can be captured—in various contexts—by detailed fluctuation theorems 2.4. For instance, the Crooks fluctuation theorem quantifies the amount of dissipation associated with a particular time-asymmetric dynamics through the ratio of trajectory probabilities (2.43).

For the discrete control-parameter dynamics considered in this chapter, we can quantify the dissipation associated with a particular transition $\boldsymbol{\lambda}_i \leftrightarrow \boldsymbol{\lambda}_j$ through the log-ratio of transition rates with the generalized detailed-balance relationship

$$\beta W^{\mathrm{SB}}_{\mathrm{ex},ji} = \ln \frac{V_{ji}}{V_{ij}} \ . \tag{8.12}$$

Here $\beta W^{\mathrm{SB}}_{\mathrm{ex},ij}$ is the 'symmetry-breaking' excess work associated with generating asymmetric jump dynamics. At the trajectory level, this relation implies that, when a particular transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ occurs, we can attribute a dissipation of $\beta W^{\mathrm{SB}}_{\mathrm{ex},ji}$, whereas the reverse process $\boldsymbol{\lambda}_j \to \boldsymbol{\lambda}_i$ results in a negative contribution $\beta W^{\mathrm{SB}}_{\mathrm{ex},ij} = -\beta W^{\mathrm{SB}}_{\mathrm{ex},ji}$.

Thus, for a one-dimensional fixed-boundary protocol ensemble—where all protocols travel between the same initial and final states $\lambda_0 \to \lambda_{N_\Lambda}$—the dissipative cost of control can be quantified by simply summing $\beta W^{\mathrm{SB}}_{\mathrm{ex},ji}$ over the net number of forward steps N_{Fw}. For fixed N_{Fw} and uniform forward (reverse) transition rates V_+ (V_-), the total dissipation of the protocol ensemble due to symmetry breaking in the protocol dynamics is therefore

$$\beta W^{\mathrm{SB}}_{\mathrm{ex},\Omega} = N_{\mathrm{Fw}} \ln \frac{V_+}{V_-} \ . \tag{8.13}$$

Combining (8.5b) with (8.13) gives a general near-equilibrium estimate of the excess work in a microscopic system driven by discrete, Markovian control-parameter dynamics

$$\langle \beta W_{\mathrm{ex}} \rangle^{\Omega}_{ji} = \langle \beta W^{\mathrm{SB}}_{\mathrm{ex}} \rangle_\Omega + \langle \beta W^{\infty}_{\mathrm{ex}} \rangle_\Omega + \langle \beta W^{\mathrm{neq}}_{\mathrm{ex}} \rangle_\Omega \tag{8.14a}$$

$$\approx \ln \frac{V_{ji}}{V_{ij}} + \beta \Delta\lambda^k_{ji} \Delta\lambda^{k'}_{ji} \left[\frac{\beta}{2} \langle \delta f_k \delta f_{k'} \rangle_{\boldsymbol{\lambda}_i} + \frac{V_{ji}}{\tilde{V}^{\mathrm{tot}}_i} \zeta^{(\alpha)}_{kk'}(\boldsymbol{\lambda}_i) \sum_s V_{is} \frac{\Delta\lambda^{k'}_{is}}{\Delta\lambda^k_{ji}} \right] \ , \tag{8.14b}$$

which now accounts for the near-equilibrium excess work—incorporating both the infinite-time and leading-order nonequilibrium contributions—as well as the dissipative cost in generating asymmetric driving dynamics.

8.4 Harmonic system

We illustrate our theoretical predictions in the simple model system of a Brownian particle diffusing in a one-dimensional quadratic potential (2.7.1),

$$E(x|\lambda_i) = \frac{k_{\text{trap}}}{2}(x - \lambda_i)^2 \ , \tag{8.15}$$

where the energetic minimum is given by the current control parameter value λ_i, and the trap stiffness is k_{trap}. The control parameter λ undergoes a discrete hopping dynamics on an infinite one-dimensional lattice of states, all separated from one another by a constant shift $\Delta\lambda$. Furthermore, we assume uniform jump rates $V_\pm$ corresponding to the forward $(+)$ and reverse $(-)$ jump rates.

Here, for exponential dwell-times, the generalized friction is equal to the viscous friction $\zeta = \gamma$ and independent of the control-parameter value. The same cannot be said, however, for non-exponential dwell times. For Gamma-distributed dwell times (8.6), the non-exponential generalized friction (8.7) can be calculated analytically (see Appendix E.2):

$$\zeta^{(\alpha)} = k_{\text{trap}}^{1-\alpha}\gamma^\alpha \ . \tag{8.16}$$

Here, the total excess work associated with a forward $(+)$ or reverse $(-)$ transition is (8.14)

$$\langle \beta W_{\text{ex}} \rangle_\pm = \ln \frac{V_\pm}{V_\mp} + \frac{\beta k_{\text{trap}} \Delta\lambda^2}{2} + \frac{\beta V_\pm \zeta^{(\alpha)} \Delta\lambda^2}{(V_+ + V_-)}(V_\pm - V_\mp) \ . \tag{8.17}$$

Thus, for a fixed-displacement control-protocol ensemble, where each protocol begins at λ_0 and terminates as soon as the control parameter arrives at λ_N, we can calculate the protocol-ensemble averaged excess work by adding the contributions associated with all of the forward and reverse steps. Separately, the three contributions to the excess work are

$$\langle \beta W_{\text{ex}}^{\text{SB}} \rangle_\Omega = \langle N_+ \rangle_\Omega \ln \frac{V_+}{V_-} + \langle N_- \rangle_\Omega \ln \frac{V_-}{V_+} \tag{8.18a}$$

$$= N_{\text{Fw}} \ln \mathcal{R} \tag{8.18b}$$

$$\langle \beta W_{\text{ex}}^{\infty} \rangle_\Omega = \frac{\beta k_{\text{trap}} \Delta\lambda^2}{2} \langle N_{\text{tot}} \rangle_\Omega \tag{8.18c}$$

$$\langle \beta W_{\text{ex}}^{\text{neq}} \rangle_\Omega = \frac{\beta \zeta^{(\alpha)} \Delta\lambda^2}{(V_+ + V_-)}(V_+ - V_-)\left[\langle N_+ \rangle_\Omega V_+ - \langle N_- \rangle_\Omega V_-\right] \ , \tag{8.18d}$$

where $\mathcal{R} \equiv V_+/V_-$ is the control-parameter transition-rate ratio, $N_{\mathrm{Fw}} \equiv \langle N_+ \rangle_\Omega - \langle N_- \rangle_\Omega$ is the net number of forward (+-directed) steps, and $\langle N_{\mathrm{tot}} \rangle_\Omega \equiv \langle N_+ \rangle_\Omega + \langle N_- \rangle_\Omega$ is the average total number of steps taken. Given that we are considering fixed-distance protocol ensembles, the net number of forward steps is fixed by the boundary conditions. As the protocol ensemble approaches a deterministic limit, where $\mathcal{R} \to \infty$, the symmetry-breaking contribution diverges [180]. [2]

We can further simplify the infinite-time excess work by calculating an explicit form for the protocol-ensemble average number $\langle N_{\mathrm{tot}} \rangle_\Omega = N_{\mathrm{Fw}}(V_+ + V_-)/(V_+ - V_-)$ of steps for uniform rates (see Appendix E.3 for details) to give

$$\langle \beta W_{\mathrm{ex}}^{\infty} \rangle_\Omega = \frac{N_{\mathrm{Fw}} \beta k_{\mathrm{trap}} \Delta\lambda^2}{2} \frac{V_+ + V_-}{V_+ - V_-} \tag{8.19a}$$

$$= \frac{N_{\mathrm{Fw}} \beta k_{\mathrm{trap}} \Delta\lambda^2}{2} \frac{\mathcal{R} + 1}{\mathcal{R} - 1} \tag{8.19b}$$

which diverges as the forward/reverse rates approach one another ($\mathcal{R} \to 1$).

The nonequilibrium excess work (8.18d) can also be simplified by writing the ensemble averages of the forward and reverse step numbers $\langle N_\pm \rangle_\Omega$ in terms of the net number of forward steps N_{Fw}, and the average total number of steps $\langle N_{\mathrm{tot}} \rangle_\Omega$:

$$\langle N_\pm \rangle_\Omega = \frac{\langle N_{\mathrm{tot}} \rangle_\Omega \pm N_{\mathrm{Fw}}}{2} \, . \tag{8.20}$$

Thus, using (8.20), the nonequilbrium excess work (8.18d) is

$$\langle \beta W_{\mathrm{ex}}^{\mathrm{neq}} \rangle_\Omega = \beta\zeta \Delta\lambda^2 N_{\mathrm{Fw}} V_+ \frac{V_+ - V_-}{V_+ + V_-} \tag{8.21a}$$

$$= \beta\zeta \Delta\lambda^2 N_{\mathrm{Fw}} V_+ \frac{\mathcal{R} - 1}{\mathcal{R} + 1} \, . \tag{8.21b}$$

In contrast to the symmetry-breaking work (8.18b) and the infinite-time excess work (8.19b), the nonequilibrium excess work does not diverge in any limit, but approaches zero for symmetric dynamics (as $\mathcal{R} \to 1$) and asymptotes to a constant value in the deterministic limit (as $\mathcal{R} \to \infty$).

8.4.1 Timescale-separated limit

In the infinite-time limit, the timescale of control-parameter dynamics is much slower than the system relaxation: $V_\pm \zeta \ll 1$. Here, the nonequilibrium excess work (8.21b) is negligible, and the protocol-ensemble averaged excess work is thus given by the sum of the symmetry-

[2] Here, we refer to the $\mathcal{R} \to \infty$ limit as a 'deterministic limit' because the probability of a reverse step becomes vanishingly small, and thus there is no randomness to the dynamics.

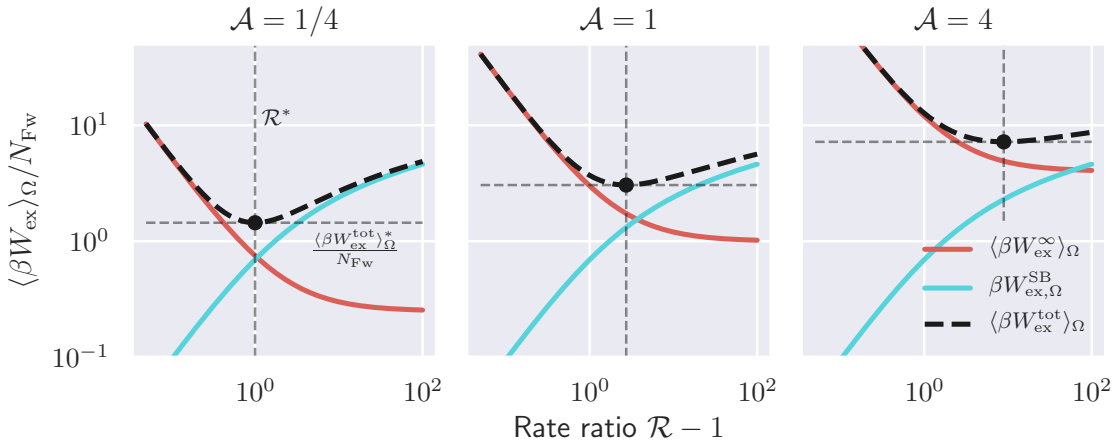

Figure 8.1: **Dissipation bound in the infinite-time limit.** In the timescale-separated limit, where the system comes to equilibrium between consecutive control-parameter jumps, the dissipation lower bound in the total work (black dashed) (8.22) arises due to a competition between the infinite-time work (red) (8.19b) and the dissipation arising due to symmetry breaking (blue) (8.18a) in the dynamics of the control parameter. Vertical and horizontal dashed lines respectively indicate the optimal rate ratio $\mathcal{R}^* - 1$ and lower bound on the excess work per forward step $\langle \beta W_{\rm ex} \rangle_\Omega^* / N_{\rm Fw}$, and intersect at the minimum (black dot). Here, $\mathcal{A} \equiv \frac{\beta}{2} k_{\rm trap} \Delta\lambda^2$ is the (unitless) prefactor of the infinite-time work (8.19b).

breaking (8.18b) and infinite-time terms (8.19b),

$$\langle \beta W_{\rm ex} \rangle_\Omega = N_{\rm Fw} \left[\ln \mathcal{R} + \frac{\beta k_{\rm trap} \Delta\lambda^2}{2} \frac{\mathcal{R}+1}{\mathcal{R}-1} \right] . \tag{8.22}$$

The total excess work in (8.22) can be minimized explicitly with respect to $\mathcal{R}$—holding all else constant—by extremizing the partial derivative

$$0 = \partial_{\mathcal{R}} \langle \beta W_{\rm ex} \rangle_\Omega \tag{8.23a}$$

$$= \frac{1}{\mathcal{R}} + \mathcal{A} \frac{2}{(\mathcal{R}-1)^2} , \tag{8.23b}$$

where $\mathcal{A} \equiv \frac{\beta}{2} k_{\rm trap} \Delta\lambda^2 > 0$.

Here, (8.23) is minimized for the optimal rate ratio $\mathcal{R}^* = \mathcal{A} + \sqrt{\mathcal{A}(\mathcal{A}+2)} + 1$, which is independent of the protocol distance $N_{\rm Fw}$. Figure 8.1 shows the infinite-time excess work (8.22) as a function of the rate ratio $\mathcal{R} - 1$ for $\mathcal{A} = 1/2, 1, 4$. In each case, there is a lower bound on the protocol-ensemble averaged excess work at a finite $\mathcal{R}^*$ due to the competition between the two terms contributing to the excess work.

8.4.2 Nonequilibrium excess work

Beyond the timescale-separated limit, where the system microstates do not fully equilibrate between discrete jumps, the nonequilibrium contribution to the total dissipation must also be included in analysis. In the absence of a control cost, the dissipation associated with a nonequilibrium protocol is given by the sum of the infinite-time and nonequilibrium contribution, which for this simple system is

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Omega} = N_{\mathrm{Fw}} \beta \Delta\lambda^2 \left[\frac{k_{\mathrm{trap}}}{2} \frac{\mathcal{R}+1}{\mathcal{R}-1} + \zeta^{(\alpha)} V_+ \frac{\mathcal{R}-1}{\mathcal{R}+1} \right] . \tag{8.24}$$

Again, we can minimize (8.24) with respect to $\mathcal{R}$ (keeping V_+ constant) to find two distinct solutions. For the ratio $\mathcal{B} \equiv 2\zeta^{(\alpha)} V_+ / k_{\mathrm{trap}}$, if $\mathcal{B} > 1$—indicating that the prefactor $\zeta^{(\alpha)} V_+$ of the nonequilibrium excess work is larger than the prefactor $k_{\mathrm{trap}}/2$ of the infinite-time work—the optimal rate ratio occurs at

$$\mathcal{R}^* = \frac{2}{\sqrt{\mathcal{B}} - 1} + 1 , \tag{8.25}$$

which is non-negative. However, for $\mathcal{B} \leq 1$ there is no finite $\mathcal{R}$ value that minimizes the excess work. Rather, the excess work is minimized asymptotically in the deterministic limit where $\mathcal{R} \to \infty$. Here, the excess work approaches the constant value

$$\lim_{\mathcal{R}\to\infty} \langle \beta W_{\mathrm{ex}} \rangle_{\Omega} = \frac{N_{\mathrm{Fw}}\, \beta k_{\mathrm{trap}} \Delta\lambda^2}{2} (1 + \mathcal{B}) , \tag{8.26}$$

implying that deterministic dynamics gives the most efficient protocol ensemble. Intuitively, this transition occurs due to both the infinite-time and nonequilibrium excess-work contributions saturating at finite values for $\mathcal{R} \to \infty$, so that their sum also approaches a constant value in the same limit (given by (8.26)). If $\langle \beta W_{\mathrm{ex}}^{\infty} \rangle > \langle \beta W_{\mathrm{ex}}^{\mathrm{neq}} \rangle$ over all rate ratios $\mathcal{R}$, then the minimum of the sum is fixed at $\mathcal{R} \to \infty$, while for $\mathcal{B} > 1$ there is a minimum for finite $\mathcal{B}$. Figure 8.2 shows (for $\alpha = 1$) the total excess work from (8.24) for $\mathcal{B} < 1$ (left), $\mathcal{B} = 1$ (middle), and $\mathcal{B} > 1$ (right), which show no finite-$\mathcal{R}$ minimum ($\mathcal{B} \leq 1$) or one finite-$\mathcal{R}$ minimum ($\mathcal{B} > 1$).

Alternatively, to compare these results with those in Ch. 6—where, in a similar system, a lower bound on excess work was found when looking at the excess work as a function of average protocol duration—we can express the total protocol-ensemble averaged excess work in terms of the average protocol duration $\langle \tau \rangle_{\Omega}$ (instead of the rate ratio $\mathcal{R}$). Given that the average dwell time in a particular state is the inverse of the total rate $V^{\mathrm{tot}} = V_+ + V_-$, the average protocol duration is $\langle \tau \rangle_{\Lambda} = \sum_{i=0}^{N_\Lambda} 1/(V_+ + V_-) = N_\Lambda/(V_+ - V_-)$ and thus, for a

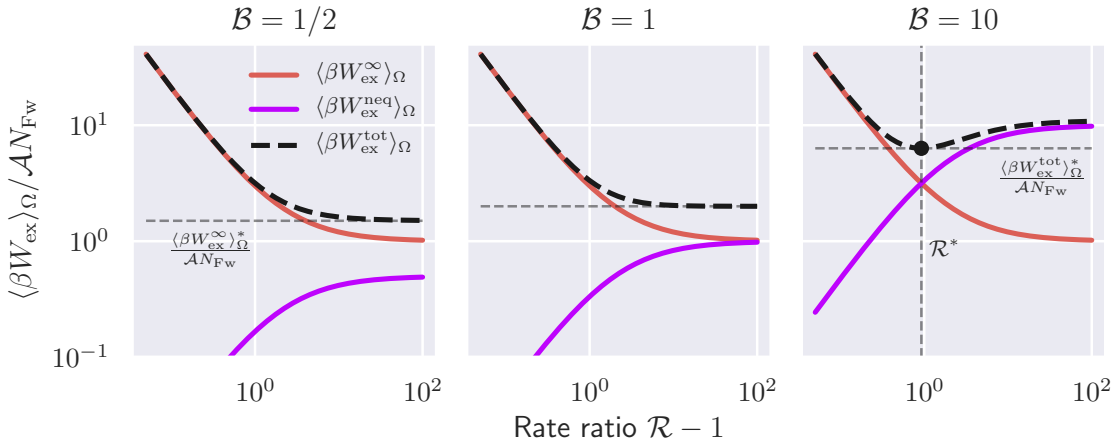

Figure 8.2: **Dissipation bound for nonequilibrium and infinite-time excess-work contributions**. Nonequilibrium (purple) and infinite-time (red) contributions to the total excess work (black dashed lines). For the excess work in (8.24), systems whose dissipation in the large-$\mathcal{R}$ limit is dominated by the nonequilibrium excess work (right) show a finite-$\mathcal{R}$ minimum of the total dissipation at an optimal $\mathcal{R}^*$. Conversely, for $\mathcal{B} \equiv 2\zeta^{(\alpha)} V_+ / k_{\rm trap}$, when $\mathcal{B} \leq 1$ there is no finite $\mathcal{R}$ which minimizes the dissipation; rather, it is minimized in the $\mathcal{R} \to \infty$ limit, where the dynamics become deterministic. In all subplots $\alpha = 1$.

fixed set of rates $V_\pm$, the protocol-ensemble averaged duration is

$$\langle \tau \rangle_\Omega = \frac{\langle N_{\rm tot} \rangle_\Omega}{V_+ + V_-} \tag{8.27a}$$

$$= \frac{N_{\rm Fw}}{V_+ - V_-} \ . \tag{8.27b}$$

To vary the average protocol duration, we can adjust the differences in the jump rates $V_+ - V_-$. Thus, writing (8.24) in terms of average protocol duration gives

$$\langle \beta W_{\rm ex} \rangle_\Omega = \beta \Delta\lambda^2 \left[\frac{k_{\rm trap}}{2} (V_+ + V_-) \langle \tau \rangle_\Omega + \frac{\zeta^{(\alpha)} N_{\rm Fw}^2}{\langle \tau \rangle_\Omega} \right] \ , \tag{8.28}$$

where the terms in brackets represent the infinite-time and nonequilibrium excess works, respectively. The scaling of each term with protocol duration is reminiscent of the protocol-ensemble averaged excess work derived in Ch. 6, with the infinite-time contribution scaling linearly and the nonequilibrium excess work scaling inversely with average protocol duration $\langle \tau \rangle_\Omega$. This competition in scaling behaviors further clarifies the presence of a lower bound on dissipation.

The protocol-ensemble averaged excess work in (8.28) can also be minimized relative to the average protocol duration, holding the total rate $(V_+ + V_-)$ fixed, achieving minimal

excess work

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Omega}^{*} = \beta N_{\mathrm{Fw}} \Delta\lambda^2 \sqrt{2 k_{\mathrm{trap}} \zeta^{(\alpha)} (V_+ + V_-)} \tag{8.29}$$

at average duration

$$\langle \tau \rangle_{\Omega}^{*} = N_{\mathrm{Fw}} \sqrt{\frac{2\zeta^{(\alpha)}}{k_{\mathrm{trap}}(V_+ + V_-)}} \, . \tag{8.30}$$

8.4.3 General dissipation bound

In general, all three contributions to the excess work will enter the total protocol-ensemble averaged excess work. Thus by combining (8.22) and (8.24) we arrive at the general expression for excess work in the harmonic system

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Omega} = N_{\mathrm{Fw}} \left[\ln \mathcal{R} + \frac{\beta k_{\mathrm{trap}} \Delta\lambda^2}{2} \frac{\mathcal{R}+1}{\mathcal{R}-1} + \beta \Delta\lambda^2 \zeta^{(\alpha)} V_+ \frac{\mathcal{R}-1}{\mathcal{R}+1} \right] \, , \tag{8.31}$$

which is given in terms of the dimensionless parameters $\mathcal{A} \equiv \beta k_{\mathrm{trap}} \Delta\lambda^2/2$ and $\mathcal{B} \equiv 2\zeta^{(\alpha)} V_+ / k_{\mathrm{trap}}$ as

$$\langle \beta W_{\mathrm{ex}} \rangle_{\Omega} = N_{\mathrm{Fw}} \left[\ln \mathcal{R} + \mathcal{A} \frac{\mathcal{R}+1}{\mathcal{R}-1} + \mathcal{A}\mathcal{B} \frac{\mathcal{R}-1}{\mathcal{R}+1} \right] \, . \tag{8.32}$$

If $\mathcal{B} \ll 1$, then the excess work function is well approximated by the timescale-separated result in 8.4.1, while when $\mathcal{A} \gg 1$ the excess work is well approximated by the excess-work function in 8.4.2. When $\mathcal{A}, \mathcal{B} \approx 1$, the excess work is a combination of the infinite-time, nonequilibrium, and symmetry-breaking contributions.

While the general expression for the optimal ratio $\mathcal{R}$ (8.32) is not particularly illuminating, the excess work is in all cases lower bounded. Figure 8.3 shows numerical results for the full dissipation (8.32) as a function of the rate ratio $\mathcal{R} - 1$ for a range of $\mathcal{A} = 1/4, 1, 4$ and $\mathcal{B} = 1/2, 1, 10$. The top row ($\mathcal{B} = 1/4$) is roughly equivalent to the result in Sec. 8.4.1 (and those shown in Fig. 8.1) where the nonequilibrium excess work (8.5b) is negligible in comparison to the other contributions. Conversely, in the bottom-right plot ($\mathcal{B} = 10$, $\mathcal{A} = 4$) the infinite-time contribution is negligible when compared to the nonequilibrium excess work (8.5b) (in the region near the minimal dissipation), similar to the results in Sec. 8.4.2 (and shown in the right subplot of Fig. 8.2) where the symmetry-breaking contribution to the excess work is not included in the analysis.

8.5 Discussion

In this chapter, we have derived an excess-work cost function for near-equilibrium systems driven by a discrete stochastic control parameter. By assuming that the control-parameter dynamics are independent of the system state, we could directly generalize the results for discrete protocols in Ch. 7 to ensembles of control protocols. Furthermore, by incorporating the dissipation due to breaking time-reversal symmetry of the control-parameter dynamics

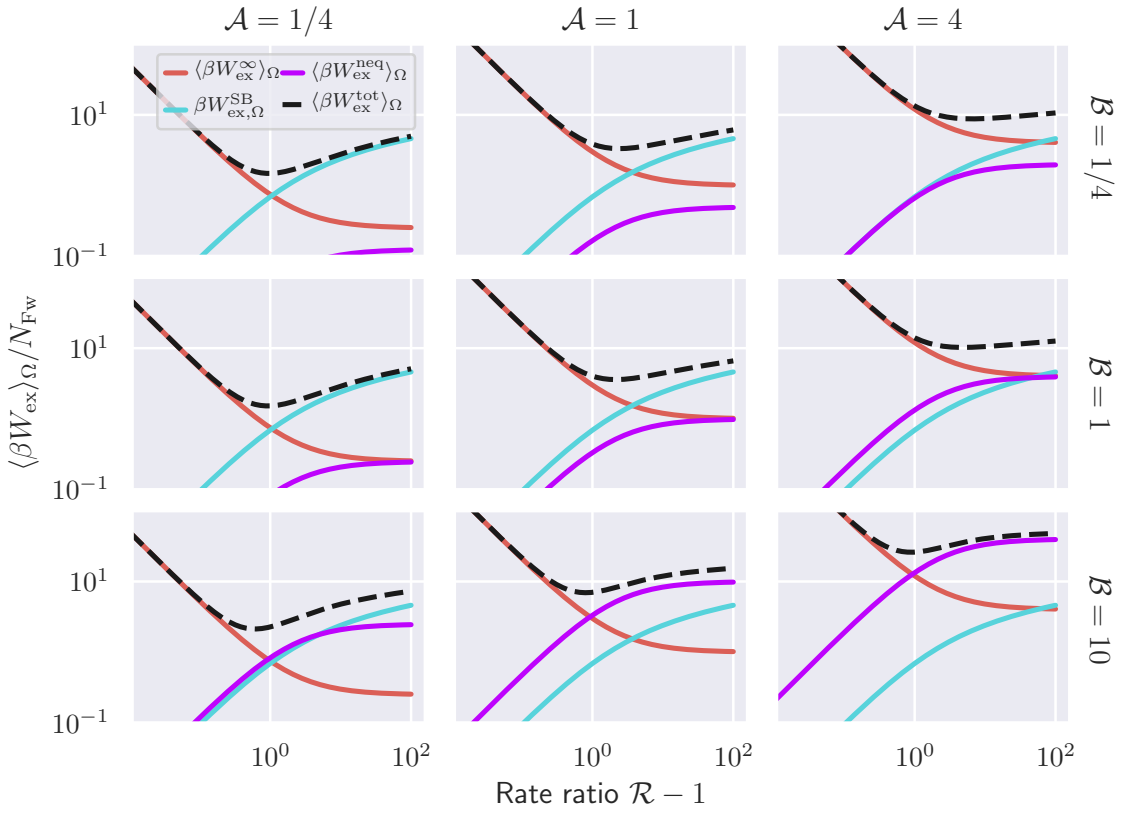

Figure 8.3: **General dissipation bound for near-equilibrium driving processes.** The excess work in (8.31) is composed of three contributions: the infinite-time work (red, Eq. (8.19b)), the nonequilibrium excess work (purple, Eq. (8.21b)), and the symmetry-breaking work (teal, Eq. 8.4.1). In all subplots $\alpha = 1$.

we connect with other results from the literature [180, 190]. We find that, although the lower bounds in Ch. 6 and [180, 190] appear mathematically similar, they arise due to different physical phenomena. In both cases, the presence of a lower bound is due to a contribution to the total dissipation that is positive for any discrete jump. We also derived a leading-order correction to the timescale-separated work that accounts for the system being out of equilibrium, found by integrating the time-dependent term in the total excess work over the distribution of dwell times [77]. We derived the nonequilibrium excess work for an exponential dwell-time distribution—in which case the generalized friction tensor [109] determines the dissipation—as well as Gamma-distributed dwell times, for which we derived a modified form of the generalized friction tensor (8.7).

In the timescale-separated limit, where the system very nearly equilibrates after each control-parameter transition, we find the lower bound on the dissipation is captured by the considerations from Ref.[180]. Alternatively, when the nonequilibrium excess work becomes significant and overcomes the symmetry-breaking dissipation, the dissipation bound is well

approximated by neglecting the asymmetry in driving dynamics and is thus similar to the bound derived in Ch. 6, which arose due to a trade-off between a continually accumulating contribution to the excess work, and a contribution which contributes more when the system is farther from equilibrium (6.26). Furthermore, it was recognized in [180, 190] that the lower bound appearing in the infinite-time limit scales subextensively with system size, as the lower bound was equal (up to a factor of 2) to the thermodynamic length $\mathcal{L}$ as measured with respect to the Fisher information matrix (FIM) [124]. $\mathcal{L}$ scales as the square root of the FIM, which effectively means that, as the system size increases, the symmetry-breaking contribution to the dissipation (8.18b) increases slower than the infinite-time contribution, and thus for large systems the symmetry-breaking contribution to the excess work is negligible. However, by including the nonequilibrium excess work, which scales with system size in the same way as the infinite-time work, the lower bound (if present) will persist. Thus, macroscopic systems which proceed through a series of discrete jumps exhibit a minimal dissipation which is determined by the trade-offs embedded in the infinite-time and nonequilibrium excess works (7.16).

While these results have potentially interesting physical implications for driven, microscopic systems, their utility in fully autonomous systems—such as molecular machines–is limited. The fundamental feature of the excess work that leads to such lower bounds is the presence of a contribution to the dissipation (here infinite-time excess work (7.16)) that is positive for any control-parameter step. However, for truly autonomous systems, the transition rates must obey a local detailed-balance relationship (2.20), and thus the assumption of control-parameter transition rates that are independent of the microscopic state of the system is problematic. In deriving the infinite-time excess-work contribution in [77, 109], it was assumed that a control-parameter transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ occurs independently of the state of the system x, and thus when the transition occurs, the system is in its equilibrium distribution (1.1). For systems which obey local detailed balance—even in the timescale-separated limit—the distribution of states from which transitions occur (known as the switching-state distribution in Ch. 9) is not the equilibrium distribution.

Put differently, the derivations in this part of the thesis (chapters 6– 8) assume that the control parameter receives no feedback from the system. However, the local detailed-balance condition (2.20) on the transition rates of a truly autonomous system require a particular type of feedback. In the following two chapters (Ch. 9 and 10), we detail the specific role that excess work plays in the dynamics of strongly coupled stochastic systems which obey local detailed balance, and thus truly represent the physics of molecular machines.

However, the derivations in this chapter—and those preceding it—may play an important role yet in the physics of more exotic systems, such as active matter. The burgeoning field of active matter is concerned with the physics of (often microscopic) agents which are not constrained in the same ways by the laws of microscopic detailed balance, and thus may be subjected to similar optimization trade-offs as outlined in the preceding chapters.

Part III

The nonequilibrium physics of autonomous machines

Chapter 9

Free energy transduction within autonomous systems

9.1 Introduction

In Part II of this thesis, we discussed several generalizations of the linear-response approach to quantifying excess work in driven nonequilibrium systems. Ultimately, we were motivated by a better understanding of molecular machines. However, we also discussed in Chapter 8 that such approximations do not incorporate feedback from the system being driven on the dynamics of the control parameter. Thus, it remains to be clarified how such control-theoretical approaches can be used to understand the physics of energy flows between the components of thermodynamic systems that obey the necessary constraints on their dynamics. Furthermore, this begs the question as to what role the internal flows of energy play in the physics of autonomous systems more generally.

This Chapter aims to elaborate on these questions, by quantifying the flow of energy between the components of a stochastic multi-component system and rigorously show how it relates to the entropy production. Ultimately, entropy production represents a fundamental loss in the operation of a machine: low-entropy states of a system (all gas molecules confined to one half of a cylinder, a fully extended polymer) or of thermal baths (baths at different temperatures) can be harnessed to do useful work (push a piston during expansion, exert a force during compaction, drive a heat engine during heat flow), whereas high-entropy states have reduced capacity to do so [194]. Thus in biophysical molecular machines, entropy production serves as an important performance measure.

Within the typical experimental paradigm, where a stochastic system is manipulated using an external device, there is a convenient correspondence between the excess work done on the system and the total entropy production [97]. Specifically, the entropy production fluctuation theorem [12] and the Crooks fluctuation theorem [7] are equivalent for a stochastic system—initially at thermodynamic equilibrium—driven by an external control parameter that receives no feedback from the system [97]. Here, the excess work and entropy

production can be used interchangeably when referring to dissipation. This correspondence is a powerful tool, granting the freedom to quantify dissipation either through work done in excess of free energy changes, or system entropy changes not compensated by heat flows. However, the same such correspondence does not hold in more complex coupled systems, such as molecular machines [195].

Although such machines operate autonomously, there are many natural comparisons between controlled systems and molecular machines. For instance, it is conceptually straightforward to view the chemical hydrolysis of ATP as driving the processive motion of the molecular motor kinesin [196], or the F_o component of ATP synthase as mechanically driving the rotation of a central crankshaft, inducing the F_1 component to catalyze the production of ATP [197]. Thus there is a natural appeal to quantifying the work (and the excess work) done by one component of an autonomous biomolecular machine on another. Indeed, such internal energy flows have been used to aid in the identification of reaction coordinates in biomolecular dynamics [198].

Given the biophysical importance of entropy production, and the mathematical constraints imposed on it by fluctuation theorems and second-law-like inequalities, it is tempting to use the correspondence between excess work and entropy production in systems driven by an external control parameter to aid in the conceptual and quantitative understanding of biomolecular machines [76, 180, 190]. However, while it is still possible to define excess work internal to such systems as an energy flow, it has no direct relationship to the entropy production.

In this chapter, we investigate excess work and entropy production in strongly coupled autonomous systems, and present an alternative measure of dissipation—the *transduced additional free energy rate*—which plays the same thermodynamic role in autonomous systems as excess power does in externally driven systems. Specifically, the steady-state transduced additional free energy rate (differing from the excess power by an information rate that captures the effects of inter-system coupling) equals an entropy production rate. We then illustrate our results in a simple model of a cyclic mechanochemical motor, where a mechanical system is driven via its strong coupling to a stochastic nonequilibrium chemical reaction.

9.2 Strongly coupled multi-component systems

Throughout this chapter, we consider a bipartite system described by two coordinates X and Y, with states x and y, respectively. The system's joint dynamics are described by a discrete-state master equation [195, 199]:

$$\mathrm{d}_t p_{xy} = \sum_{x',y'} R^{xx'}_{yy'} p_{x'y'} \ , \tag{9.1}$$

where p_{xy} is the joint probability of state (x, y), $\mathrm{d}_t \equiv \mathrm{d}/\mathrm{d}t$ is the time derivative, and $R^{xx'}_{yy'}$ is the transition rate matrix element quantifying the rate of the transition $(x', y') \to (x, y)$. Such systems are said to be strongly coupled, as the energy of a particular state x (y) depends on the state of Y (X).

The assumption of bipartite dynamics restricts the entries of the transition rate matrix [195, 199]:

$$R^{xx'}_{yy'} = \begin{cases} R^{x}_{yy'} & \text{if} \quad y \neq y' \quad \text{and} \quad x = x' \\ R^{xx'}_{y} & \text{if} \quad y = y' \quad \text{and} \quad x \neq x' \\ 0 & \text{otherwise} \end{cases} . \tag{9.2}$$

Thus, for the ensemble average $\langle A \rangle_{XY} \equiv \sum_{x,y} p_{xy} a_{xy}$ of the fluctuating quantity a_{xy} that depends on the joint state (x, y), the time rate of change $\mathrm{d}_t \langle A \rangle_{XY} = \sum_{x,x',y,y'} R^{xx'}_{yy'} p_{x'y'} (a_{xy} - a_{x'y'})$ can be decomposed into contributions due to the individual dynamics of each subsystem:

$$\mathrm{d}_t \langle A \rangle_{XY} = \sum_{x,x',y} R^{xx'}_{y} p_{x'y} (a_{xy} - a_{x'y}) + \sum_{x,y,y'} R^{x}_{yy'} p_{xy'} (a_{xy} - a_{xy'}) \tag{9.3a}$$

$$= \langle \dot{A} \rangle^X + \langle \dot{A} \rangle^Y . \tag{9.3b}$$

The overdot indicates that—unlike the time derivative $\mathrm{d}_t \langle A \rangle_{XY}$—the individual rates ($\langle \dot{A} \rangle^X$ and $\langle \dot{A} \rangle^Y$) cannot be written as the time derivative of a function. Here, angle brackets $\langle \cdots \rangle_{XY}$ indicate an average over the joint distribution p_{xy}, while $\langle \cdots \rangle^X$ ($\langle \cdots \rangle^Y$) indicate an average over the dynamics of the X (Y) subsystem alone.

9.2.1 Entropy production

The entropy production rate (for unitless entropy) of the joint system is [195]

$$\dot{\Sigma} = \sum_{x,y,x',y'} R^{xx'}_{yy'} p_{x'y'} \ln \frac{R^{xx'}_{yy'} p_{x'y'}}{R^{x'x}_{y'y} p_{xy}} \geq 0 . \tag{9.4}$$

The total entropy production [1] can be conveniently split into separate contributions arising from the respective transitions among subsystems Y and X:

$$\dot{\Sigma} = \dot{\Sigma}^Y + \dot{\Sigma}^X \tag{9.5a}$$

$$= \sum_{x,y,y'} R^{x}_{yy'} p_{xy'} \ln \frac{R^{x}_{yy'} p_{xy'}}{R^{x}_{y'y} p_{xy}} + \sum_{x,x',y} R^{xx'}_{y} p_{x'y} \ln \frac{R^{xx'}_{y} p_{x'y}}{R^{x'x}_{y} p_{xy}} . \tag{9.5b}$$

[1] Because the entropy S (Sec. 2.5) is defined as the ensemble average of $-\ln p_x$, we do not use angle brackets around entropy and entropy production terms here.

Each RHS entropy production rate obeys a second-law-like inequality,

$$\dot{\Sigma}^X = \mathrm{d}_t S_X + \dot{S}_\mathrm{e}^X - \dot{I}^X \geq 0 \tag{9.6a}$$

$$\dot{\Sigma}^Y = \mathrm{d}_t S_Y + \dot{S}_\mathrm{e}^Y - \dot{I}^Y \geq 0 \ . \tag{9.6b}$$

$\mathrm{d}_t S_X$ ($\mathrm{d}_t S_Y$) is the rate of change of the entropy of subsystem X (Y). $\dot{S}_\mathrm{e}^X$ ($\dot{S}_\mathrm{e}^Y$) is the rate of entropy flow from the system to the heat reservoir due to X (Y) dynamics, which for a thermodynamic system coupled to a heat bath (all systems considered here) equals $-\beta\dot{\mathcal{Q}}^X$ ($-\beta\dot{\mathcal{Q}}^Y$), the negative rate of heat flow into the system—and hence entropy increase in the heat reservoir—due to X (Y) dynamics.

The information rates are

$$\dot{I}^X = \sum_{x,x',y} R_y^{xx'} p_{x'y} \ln \frac{p_{y|x}}{p_{y|x'}} \tag{9.7a}$$

$$\dot{I}^Y = \sum_{x,y,y'} R_{yy'}^{x} p_{xy'} \ln \frac{p_{x|y}}{p_{x|y'}} \ , \tag{9.7b}$$

for conditional probability $p_{x|y} \equiv p_{xy} / \sum_x p_{xy}$ of state x given state y. $\dot{I}^X$ ($\dot{I}^Y$) represents the rate of change of mutual information between subsystems X and Y due to transitions in X (Y) [195, 199]. Thus, a positive (negative) information rate $\dot{I}^X$ implies that, on average, dynamics of X increase (decrease) the mutual information between the two subsystems (Sec. 2.5).

At steady state, the joint-system entropy is unchanging ($\mathrm{d}_t S_{XY} = 0$), as are the entropies of each subsystem ($\mathrm{d}_t S_X = \mathrm{d}_t S_Y = 0$). Thus, the combined effect of X and Y dynamics leaves the mutual information unchanged, so the information rates are opposite ($\dot{I}^X = -\dot{I}^Y$) and cancel when summing the entropy production rates $\dot{\Sigma}^X$ and $\dot{\Sigma}^Y$ of each subsystem [195] [Eq. (9.6)]. Therefore, the total steady-state entropy production is the sum of each subsystem's heat flow:

$$\dot{\Sigma} = -\langle \beta\dot{\mathcal{Q}} \rangle^Y - \langle \beta\dot{\mathcal{Q}} \rangle^X \ . \tag{9.8}$$

However, unlike the entropy productions of each subsystem (9.6), neither RHS term is lower bounded by zero.

9.2.2 Excess work

To quantify the energy flow in such a system, we now treat subsystem Y as a work source for subsystem X, so that the average rate of work (the average power) done by Y on X is

$$\langle \beta\mathcal{P} \rangle_{Y \to X} = \sum_{x,y,y'} \beta \left(\epsilon_{xy} - \epsilon_{xy'} \right) R_{yy'}^{x} p_{xy'} \ , \tag{9.9}$$

for energy ϵ_{xy} of state (x, y). Throughout this chapter we exclusively deal with average power and average work, so for the remainder we omit explicit mention of averages.

For fixed y, the conditional equilibrium distribution of X is $\pi_{x|y} = \exp(-\beta\epsilon_{xy} + \beta F_{X|y})$, for conditional equilibrium free energy $F_{X|y} \equiv -k_{\mathrm{B}}T \ln \sum_x \exp(-\beta\epsilon_{xy})$. Thus, the energy can be expressed in terms of the conditional equilibrium distribution and free energy as

$$\epsilon_{xy} = -k_{\mathrm{B}}T \ln \pi_{x|y} + F_{X|y} \ . \tag{9.10}$$

The power done on X is the sum of the reversible power and the excess power,

$$\langle \beta\mathcal{P} \rangle_{Y\to X} = \langle \beta\mathcal{P}_{\mathrm{rev}} \rangle_{Y\to X} + \langle \beta\mathcal{P}_{\mathrm{ex}} \rangle_{Y\to X} \tag{9.11a}$$

$$= \beta \sum_{x,y,y'} R^x_{yy'} p_{xy'} (F_{X|y} - F_{X|y'}) + \sum_{x,y,y'} R^x_{yy'} p_{xy'} \ln \frac{\pi_{x|y'}}{\pi_{x|y}} \ . \tag{9.11b}$$

The excess power quantifies the rate of energy flow from subsystem Y into subsystem X which exceeds the rate of change of the conditional equilibrium free energy of X.

The excess power can be simplified by separating the X and Y summations,

$$\langle \beta\mathcal{P}_{\mathrm{ex}} \rangle_{Y\to X} = \sum_{y,y'} V_{yy'} p_{y'} \left[\sum_x \frac{R^x_{yy'} p_{x|y'}}{V_{yy'}} \ln \frac{\pi_{x|y'}}{\pi_{x|y}} \right] \tag{9.12a}$$

$$= \sum_{y,y'} V_{yy'} p_{y'} \langle \beta W_{\mathrm{ex}} \rangle_{yy'} \ , \tag{9.12b}$$

where the X summation is simply the excess work on X during transition $y' \to y$. The coarse-grained rate

$$V_{yy'} \equiv \sum_x R^x_{yy'} p_{x|y'} \tag{9.13}$$

is the rate of $y' \to y$, averaged over the conditional distribution of X [68]. $V_{yy'}$ normalizes the switching-state distribution

$$p^{\mathrm{sw},x}_{yy'} \equiv \frac{R^x_{yy'} p_{x|y'}}{V_{yy'}} \ , \tag{9.14}$$

the conditional distribution of X during a $y' \to y$ transition.

The excess work per $y' \to y$ transition can be expressed using (9.14) as the difference between the relative entropies of the switching-state distribution with the respective conditional equilibrium distributions after $(\pi_{x|y})$ and before $(\pi_{x|y'})$ the transition:

$$\langle \beta W_{\mathrm{ex}} \rangle_{yy'} = D\left(p^{\mathrm{sw},x}_{yy'} \,||\, \pi_{x|y}\right) - D\left(p^{\mathrm{sw},x}_{yy'} \,||\, \pi_{x|y'}\right) \ . \tag{9.15}$$

9.3 Classes of upstream dynamics

9.3.1 External control parameter

Insofar as it relates to entropy production, the excess power is a quantity of interest in many systems driven by an external control parameter. Many experimental manipulations of machines drive the system according to dynamics of an experimental apparatus that do not depend on the system response, i.e., with no feedback. This corresponds to the special case of the above framework where the Y dynamics are independent of the current state of X. Here, we consider stochastic Y dynamics (as in Ch. 6) [180, 190]; however, in general they could alternatively be deterministic (Ch. 7) [106, 107, 109].

Since such independent Y dynamics ensure the conditional independence of the target state y and the current mechanical state x (conditioned on the source state y'), the data-processing inequality [94] requires that the Y dynamics reduce the mutual information between subsystems: $\dot{I}^Y \leq 0$. Mathematically, for independent Y dynamics the information rate $\dot{I}^Y$ (9.7b) can be written as a negative relative entropy, $\dot{I}^Y = -\sum_{y,y'} V_{yy'}\, p_{y'} D\left(p_{x|y'} || p_{x|y}\right) \leq 0$, and thus is necessarily non-positive. Therefore, at steady state (where $\dot{I} = 0$), $\dot{I}^X \geq 0$, and the form of the second law $\langle \beta \dot{\mathcal{Q}} \rangle^X \leq 0$ holds for the heat flow $\langle \beta \dot{\mathcal{Q}} \rangle^X$ due to X dynamics. [2]

For independent Y dynamics, the switching-state distribution is the conditional distribution $p_{x|y'}$ of x given the source state y', independent of the target state y, and the excess work for transition $y' \to y$ is

$$\langle \beta W_{\text{ex}} \rangle_{yy'} = D\left(p_{x|y'} \,||\, \pi_{x|y}\right) - D\left(p_{x|y'} \,||\, \pi_{x|y'}\right) \ . \tag{9.16}$$

In the timescale-separated limit [68], where the X dynamics are much faster than the Y dynamics, the conditional distribution over X equilibrates between each Y transition, so the second RHS term is zero, reducing (9.16) to the *infinite-time excess work* (7.11) from Sec. 7.3,

$$\langle \beta W_{\text{ex}} \rangle_{yy'} = D\left(\pi_{x|y'} || \pi_{x|y}\right) \geq 0 \ , \tag{9.17}$$

which persists, even in the infinite-time limit where the conditional distribution over mechanical states $p_{x|y}$ completely relaxes to equilibrium between each chemical transition. This excess work is non-negative for any transition, and thus the excess power (9.12) is positive, even for no net Y flux, $V_{yy'} p_{y'} = V_{y'y} p_y$ (Sec. 9.4.3 gives details).[3]

[2] This insight allows some intriguing interpretations in the context of the *thermodynamics of sensing*, where a system X collects information about an external and independent stochastic variable Y. Rearranging the second law with the information rate on the RHS of (9.6) yields a refined lower bound on the steady-state dissipation for the system in terms of the *nostalgia* [200, 201] or *learning rate* [202, 203].

[3] We found similar behavior in Ch. 6, where the excess power (6.26) to drive a system through an ensemble of stochastic control protocols–independent of the system response–contains a term that is independent of the driving strength.

9.3.2 Thermodynamically complete system

In autonomous systems (such as molecular machines consisting of multiple strongly interacting components) not subject to temporal variation of an external control parameter, thermodynamic consistency requires that the entries of the transition rate matrix $R^{xx'}_{yy'}$ satisfy generalized detailed balance (2.20) [204, 205, 206],

$$\ln \frac{R^{xx'}_{yy'}}{R^{x'x}_{y'y}} = -\beta \Delta \omega^{xx'}_{yy'} \ , \tag{9.18}$$

where $\Delta\omega^{xx'}_{yy'}$ is the change in a thermodynamic potential during the transition $(x', y') \to (x, y)$. Complementary to the joint system's energetic potential ϵ_{xy} which governs its interaction with a thermal reservoir, changes in thermodynamic potential can, in addition, quantify the coupling between the joint system and other types of thermodynamic reservoirs—such as chemical reservoirs through a chemical potential [204].

We call systems *thermodynamically complete* when all rates satisfy local detailed balance (9.18). Conversely, we call systems (such as the independent Y dynamics in Sec. 9.3.1) *thermodynamically incomplete* when the transition rates violate (9.18), as some external influences are required to ensure thermodynamic consistency.

In detailed-balanced dynamics—or any dynamics where subsystem Y receives feedback from X—the excess work (9.15) associated with a particular Y transition is not lower bounded by zero, and can be negative.

We present the *transduced additional free energy rate* or TAFER (the name will become clear),

$$\beta \dot{F}^{\text{add}}_{Y\to X} \equiv \langle \beta \mathcal{P}_{\text{ex}} \rangle_{Y\to X} + \dot{I}^{Y} \ , \tag{9.19}$$

as a measure of dissipation between strongly coupled subsystems that is analogous to the excess power (9.12) in systems driven by an external control parameter. Much like the excess power, $\beta \dot{F}^{\text{add}}_{Y\to X}$ averages over Y transitions while X remains fixed. Appendix F.1 gives a detailed derivation.

Unlike the excess power, however, at steady state TAFER is lower bounded by zero: $\langle \beta \mathcal{P}_{\text{ex}} \rangle_{Y\to X} = -\langle \beta \dot{\mathcal{Q}} \rangle^{X}$ (see App. F.2) and $\dot{I}^{Y} = -\dot{I}^{X}$, so (9.19) coincides with (9.6a). In terms of the underlying probability distributions,

$$\beta \dot{F}^{\text{add}}_{Y\to X} = \sum_{y,y'} V_{yy'} p_{y'} \left[\sum_{x} p^{\text{sw},x}_{yy'} \ln \frac{\pi_{x|y'} p_{x|y}}{\pi_{x|y} p_{x|y'}} \right] \ . \tag{9.20}$$

TAFER modifies the excess power in (9.12) by the additional average of the log-ratio of nonequilibrium conditional distributions over the switching-state distribution $p^{\text{sw},x}_{yy'}$. This form clarifies that TAFER vanishes for all transitions $y' \to y$ in the timescale-separated limit, where $p_{x|y} \to \pi_{x|y}$.

Furthermore, substituting (9.14) and (9.10) into (9.20), the transduced additional free energy rate can be written as

$$\beta \dot{F}^{\mathrm{add}}_{Y\to X} = \sum_{x,y,y'} R^x_{yy'} p_{xy'} \ln \frac{\pi_{x|y'} p_{x|y}}{\pi_{x|y} p_{x|y'}} \tag{9.21}$$

$$= \sum_{x,y,y'} R^x_{yy'} p_{xy'} \left[\beta \left(\epsilon_{xy} - F_{X|y} - \epsilon_{xy'} + F_{X|y'} \right) + \ln p_{x|y} - \ln p_{x|y'} \right] . \tag{9.22}$$

We now define a *stochastic nonequilibrium free energy*

$$\beta f^{\mathrm{neq}}_{x|y} \equiv \beta \epsilon_{xy} + \ln p_{x|y} \ , \tag{9.23}$$

which is the specific quantity that, when averaged over the joint distribution p_{xy}, gives the conditional nonequilibrium free energy [207]

$$\beta F^{\mathrm{neq}}_{X|Y} \equiv \langle \beta E \rangle_{XY} - S_{X|Y} \tag{9.24}$$

where $\langle E \rangle_{XY} \equiv \sum_{x,y} \epsilon_{xy} p_{xy}$ and the conditional entropy $S_{X|Y} \equiv -\sum_{x,y} p_{xy} \ln p_{x|y}$ represent averages of the energy and conditional stochastic entropy, over the joint distribution p_{xy}.

This definition allows us to rewrite Eq. (9.22) as

$$\dot{F}^{\mathrm{add}}_{Y\to X} = \sum_{x,y,y'} R^x_{yy'} p_{xy'} \left[f^{\mathrm{neq}}_{x|y} - f^{\mathrm{neq}}_{x|y'} - \left(F_{X|y} - F_{X|y'} \right) \right]$$

$$= \dot{F}^{\mathrm{neq},Y}_{X|Y} - \mathrm{d}_t F_{X|Y} \ , \tag{9.25}$$

where $\dot{F}^{\mathrm{neq},Y}_{X|Y}$ is the change of conditional nonequilibrium free energy that is due to the Y dynamics, and $\mathrm{d}_t F_{X|Y}$ is the change in equilibrium free energy of X given Y. Here, we utilized the splitting of rates of change introduced in (9.3). Finally, note that X dynamics do not modify the average equilibrium free energy $F_{X|Y}$; thus $\mathrm{d}_t F_{X|Y}$ is due to Y dynamics only.

Thus, the transduced additional free energy rate is indeed the rate of change due to the Y dynamics of the additional free energy in X [207] (above the equilibrium free energy). This clarifies that TAFER quantifies how much the Y dynamics contribute to X being out of conditional equilibrium.

9.4 Model system

We now illustrate our theory in a minimal model of a mechanochemical molecular machine. Figure 9.1 shows a schematic.

The chemical coordinate Y evolves on a 3-state cycle, while the mechanical coordinate X contains N discrete states, evolving on a periodic energy landscape imposed by the current value of Y. Subsystem X only has transitions between nearest-neighbor states, and obeys

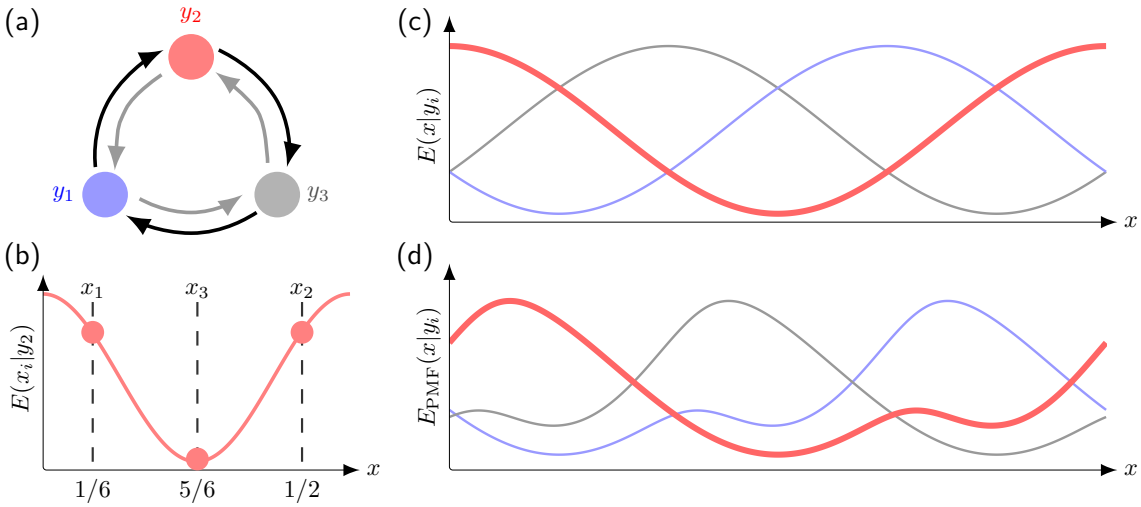

Figure 9.1: **Schematic of the model mechanochemical system.** (a) 3-state chemical reaction network representing Y dynamics. (b) The imposed energy landscape on the mechanical coordinate X (with equidistant discrete states) for $Y = y_2$ and $N = 3$ mechanical states (Sec. 9.4.1). Each specific chemical state y imposes on X a periodic energy landscape, with either (c) a single minimum (Sec. 9.4.1) or (d) two local minima (Sec. 9.4.2).

periodic boundary conditions. The instantaneous transition rates for X and Y are

$$R^x_{yy'} = \Gamma_{\mathrm{chem}} \exp\left\{-\tfrac{1}{2}\beta\left(\Delta\mu_{yy'} + \alpha\Delta\epsilon^x_{yy'}\right)\right\} \tag{9.26a}$$

$$R^{xx'}_{y} = \Gamma_{\mathrm{mech}} \exp\left\{-\tfrac{1}{2}\beta\left(\Delta\epsilon^{xx'}_{y}\right)\right\} \ , \tag{9.26b}$$

where Γ_{chem} and Γ_{mech} are kinetic prefactors for the chemical and mechanical rates, quantifying the *bare rates* of each process in the absence of any differences in state energies ϵ_{xy} or chemical potentials μ [208].[4] Here, $\alpha \in [0, 1]$ is a *feedback parameter* which interpolates the system between detailed balanced ($\alpha = 1$, Sec. 9.3.1) and feedback-free ($\alpha = 0$, Sec. 9.3.2). For all $\alpha \neq 1$ the system is thermodynamically incomplete, breaking local detailed balance (9.18), and for decreasing α the feedback from the mechanical system X to the chemical dynamics decreases.

We further assume identical chemical potential differences ($\Delta\mu_{yy'} = \Delta\mu$) for each chemical transition, and (without loss of generality) that negative $\Delta\mu$ induces net clockwise rotation of the chemical coordinate. Physically, the chemical potential differences are generated by out-of-equilibrium concentrations of products and reactants, such as ATP and ADP for many molecular machines.

[4]Unlike the energy ϵ_{xy}, the chemical potential is not a state function, and thus there is no unique chemical potential μ_y for each state y. However, changes in chemical potential upon a transition $\Delta\mu_{yy'}$ are well-defined for each transition.

In all cases considered, the steady-state excess power $\langle \beta \mathcal{P}_{\rm ex} \rangle_{Y \to X}$, TAFER $\beta \dot{F}^{\rm add}_{Y \to X}$, and entropy production rate $\dot{\Sigma}^X$ are calculated from the steady-state probability distribution. Numerically we calculate the steady-state distribution in each case by solving the eigenvalue problem outlined in App. A.1.

9.4.1 Excess power does not equal entropy production

First, we consider a thermodynamically complete system ($\alpha = 1$), where the mechanical state experiences the periodic monostable potential (Sec. 2.7.2),

$$E(x|y) = \tfrac{1}{2} E^{\ddagger} \cos 2\pi \, (x - y) \ , \tag{9.27}$$

with barrier height $E^{\ddagger}$ (see Fig. 9.1b,c). Figure 9.2 shows numerical calculations of the steady-state excess power $\langle \beta \mathcal{P}_{\rm ex} \rangle_{Y \to X}$, transduced additional free energy rate $\beta \dot{F}^{\rm add}_{Y \to X}$, and mechanical entropy production rate $\dot{\Sigma}^X$, as functions of the chemical driving strength $-\beta \Delta \mu$. For $N = 3$ mechanical states, across all barrier heights the excess power is less than the entropy production rate, while for $N = 12$, the excess power is greater. Thus, even for the simple case of $N = 3$ and X tracking the current state of Y, the excess power $\langle \beta \mathcal{P}_{\rm ex} \rangle_{Y \to X}$ can significantly differ from the entropy production rate. In contrast, the transduced additional free energy rate (9.19) equals (as expected) the entropy production rate $\dot{\Sigma}^X$ by the mechanical system's dynamics, for all energy barriers $\beta E^{\ddagger}$ and numbers N of mechanical states.

9.4.2 Excess power can become negative

At steady state (again, for $\alpha = 1$), the excess power equals the negative heat flow $-\langle \beta \dot{\mathcal{Q}} \rangle^X$ due to X dynamics, and is therefore not lower bounded by zero (see App. F.2). To illustrate this, we consider the periodic bistable potential (see Fig. 9.1d, and further discussion in Sec. 2.7.3)

$$\beta E_{\rm PMF}(x|y) = -\ln \left[e^{-\frac{1}{2}\beta E^{\ddagger} \cos 2\pi (x-y)} + e^{-\frac{1}{2}\beta E_2^{\ddagger} \cos 2\pi (x-y-\phi) + \beta \Delta E} \right] \ , \tag{9.28}$$

the potential of mean force [133] of two offset sinusoidal potentials. ϕ and ΔE represent, respectively, the relative phase shift and energy offset.

Figure 9.3 shows the steady-state transduced additional free energy rate $\beta \dot{F}^{\rm add}_{Y \to X}$, excess power $\langle \beta \mathcal{P}_{\rm ex} \rangle_{Y \to X}$, and entropy production rate $\dot{\Sigma}^X$ due to X dynamics, as a function of (a) chemical potential difference and (b) the ratio $\Gamma_{\rm mech}/\Gamma_{\rm chem}$ of bare mechanical and chemical transition rates. Once again, TAFER equals the entropy production. The excess power is negative for intermediate chemical potential differences and for large rate ratios ($\Gamma_{\rm mech}/\Gamma_{\rm chem} \gtrapprox 10^3$).

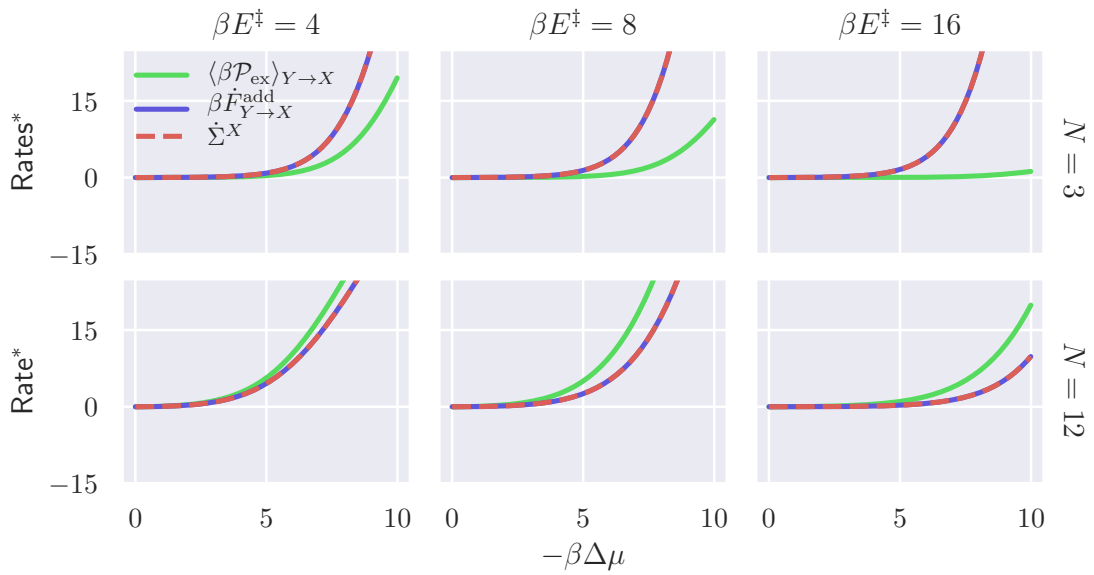

Figure 9.2: **At steady state, transduced additional free energy rate equals entropy production rate, but excess power need not.** Excess power $\langle \beta \mathcal{P}_{\mathrm{ex}} \rangle_{Y \to X}$ (green), transduced additional free energy rate $\beta \dot{F}^{\mathrm{add}}_{Y \to X}$ (blue), and entropy production rate $\dot{\Sigma}^X$ (red dashed) due to X dynamics, each as a function of the chemical driving strength $-\beta \Delta \mu$, for a range of barrier heights $\beta E^\ddagger = 4, 8, 16$ (columns) and numbers $N = 3, 12$ of mechanical states (rows). For all panels, the system is at steady state for potential (9.27), $\Gamma_{\mathrm{mech}}/\Gamma_{\mathrm{chem}} = 100$, and rates are nondimensionalized by the chemical bare rate: $\mathrm{Rate}^* \equiv \mathrm{Rate}/\Gamma_{\mathrm{chem}}$.

Physically, the excess power drops below zero because the switching-state distribution $p^{\mathrm{sw},x}_{yy'}$ for clockwise Y transitions is dominated by mechanical states x that lose energy during the transition. As a result, the excess power required to drive the system via Y dynamics becomes negative.

Thus even though the mathematical forms of the transduced additional free energy rate and the entropy production rate of the mechanical system are quite different—one involving a summation over Y transitions at fixed mechanical state x, the other involving a summation over X transitions at fixed chemical state y—they are equal at steady state. The excess power $\langle \beta \mathcal{P}_{\mathrm{ex}} \rangle_{Y \to X}$ by itself is not lower bounded by zero and can become negative (Fig. 9.3). This reinforces the mathematical demonstration (Sec. 9.3.2) that TAFER is the thermodynamic generalization of excess power to autonomous systems.

9.4.3 Entropy production in thermodynamically complete or incomplete systems

Finally, we show that, for thermodynamically incomplete systems, the entropy production $\dot{\Sigma}^X$ approaches a nonzero value even as $-\beta \Delta \mu \to 0$. To show this, we investigate the entropy production in the monostable potential (9.27) for several values of the feedback

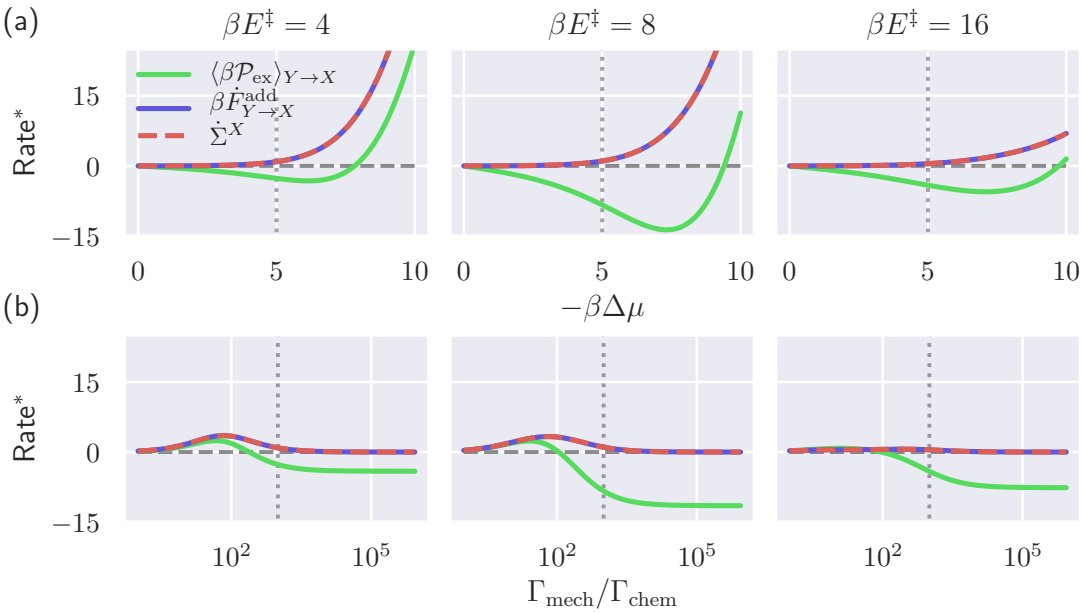

Figure 9.3: **Steady-state excess power can become negative.** Excess power $\langle \beta \mathcal{P}_{\rm ex} \rangle_{Y \to X}$ (green), transduced additional free energy rate $\beta \dot{F}^{\rm add}_{Y \to X}$ (blue), and entropy production rate $\dot{\Sigma}^X$ (red dashed), as a function (a) of chemical potential difference $-\beta \Delta \mu$ at fixed ratio $\Gamma_{\rm mech}/\Gamma_{\rm chem} = 10^4$ of bare chemical and mechanical transition rates, and (b) of the ratio $\Gamma_{\rm mech}/\Gamma_{\rm chem}$ for fixed chemical potential difference $-\beta \Delta \mu = 5$. Different columns show $\beta E^\ddagger = 4, 8, 16$. All calculations are at steady state for potential (9.28) and use $N = 12$, $E^* \equiv E^\ddagger / E^\ddagger_2 = 1$, $\phi = 1/3$, and $\Delta E / E^\ddagger = 3/8$. Rates are nondimensionalized by the chemical bare rate: Rate* $\equiv$ Rate$/\Gamma_{\rm chem}$. Dotted gray vertical lines indicate in (a) the chemical potential used in (b), and in (b) the bare-rate ratio used in (a).

parameter α. Figure 9.4 shows the entropy production rate due to subsystem X dynamics as a function of the chemical driving strength $-\beta \Delta \mu$, for energy barriers $\beta E^\ddagger = 4, 8, 16$. As $-\beta \Delta \mu \to 0$, for a thermodynamically complete system ($\alpha = 1$) the entropy production rate $\dot{\Sigma}^X$ vanishes, while for a thermodynamically incomplete system ($\alpha \neq 0$), $\dot{\Sigma}^X$ approaches a nonzero asymptotic value. The asymptotic value monotonically increases with decreasing α, and is largest for $\alpha = 0$, when Y dynamics receive no feedback from X.

9.5 Discussion

Autonomous stochastic systems are models for many molecular machines, where strong coupling between multiple stochastic coordinates is central to their functioning. It is intuitively appealing to view the interacting components of a molecular machine as driving one another.

For (non-autonomous) systems where the driver–subsystem Y in Sec. 9.4–receives no feedback from the driven subsystem (X), the excess power flowing from Y to X equals the

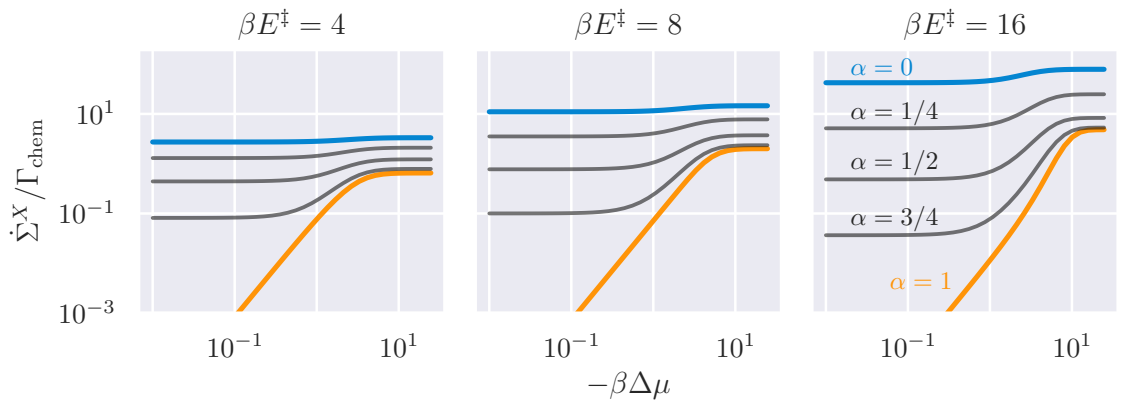

Figure 9.4: **Dynamics that break detailed balance produce nonzero steady-state entropy production, even at $-\beta\Delta\mu = 0$.** Across all energy barriers $\beta E^\ddagger = 4, 8, 16$, only detailed-balanced joint dynamics ($\alpha = 1$, orange curves) produce vanishing entropy production rate $\dot{\Sigma}^X$ as $-\beta\Delta\mu \to 0$. For all other α (blue and gray curves), the entropy production rate approaches a nonzero value as $-\beta\Delta\mu \to 0$. Y transitions that are independent of X ($\alpha = 0$, blue) have the greatest asymptotic value, while increasing feedback parameter α monotonically decreases the asymptotic entropy production rate. For all panels, the number of mechanical states is $N = 12$, and the entropy production rate is nondimensionalized by the bare chemical transition rate: $\dot{\Sigma}^X/\Gamma_{\text{chem}}$.

entropy produced by the driven subsystem's dynamics. It is tempting to make use of this correspondence between excess power and entropy production, but excess power does not play the same role in coupled, thermodynamically complete, autonomous systems and has no simple relationship with entropy production.

We introduced here the transduced additional free energy rate $\beta\dot{F}^{\text{add}}_{Y\to X}$ as a measure of dissipation in strongly coupled stochastic systems that plays the same thermodynamic role as the excess power in systems driven by an external control parameter. In particular, $\beta\dot{F}^{\text{add}}_{Y\to X}$ equals the steady-state entropy production rate during subsystem X's dynamics. Furthermore, our investigation of the transduced additional free energy rate provides a convenient generalization beyond the reversible limit of the work associated with stochastic driving protocols [183], and an analysis of the trade-offs between control work and mutual information complementary to Ref. [182].

We expect that these insights will be useful for ongoing research in stochastic thermodynamics, extending theoretical results for systems driven by an external control parameter (Sec. 6 and 7) [106, 107, 109] to autonomous models of molecular machines. A better understanding of dissipation in thermodynamically complete systems–where inter-system feedback satisfies local detailed balance–will clarify the functional capabilities and limitations of molecular machines. By consistently incorporating feedback into the control schema, we can further elucidate the rich physics in strongly coupled systems.

In the following Chapter (Ch. 10), we investigate an application of the linear-response theory approach used in part II to approximate the excess power (introduced in Sec. 9.2) in a fully autonomous system.

Chapter 10

Hidden excess power and autonomous Maxwell demons in strongly coupled nonequilibrium systems

10.1 Introduction

The previous chapter (Ch. 9) showed that, within strongly coupled systems, the thermodynamic relationship between excess power and entropy production is not the same as in externally driven (or thermodynamically incomplete) systems. However, while excess power is not, in any simple way, related to an entropy production, it still plays an important role in the physics of strongly coupled systems, such as molecular machines, for instance in identifying reaction coordinates [198]. In this chapter, we quantify the energetic flows communicated between the components in a strongly coupled mechanochemical system, near the timescale-separated (TSS) limit, using familiar tools from linear-response theory (Sec. 2.6.1) and small-perturbation expansions (App. D.1), and discuss the implications for our understanding of biological systems at the nanoscale.

Fundamentally, the consistency of stochastic thermodynamics relies on an assumption of timescale separation, allowing one to clearly distinguish the system from its surroundings [5, 68, 209]. In microscopic biological systems, however, this separation is blurred, as there are often a number of comparable timescales relevant to a given problem. As a result, it is common to model biomolecular systems as consisting of multiple strongly interacting subsystems, as seen, for instance, in F_oF_1-ATP synthase [210]. In fact, while the chemical and mechanical aspects of the motor are separable in their contributions, both are necessary to understand the dynamics of the machine [211]. This strong coupling between chemical and mechanical processes is ubiquitous among molecular machines [212].

In practice, the description of such systems is often simplified through the implicit use of a coarse-graining procedure, whereby the state space of the full mechanochemical system

is projected onto a smaller set of *mesostates*, each containing several microstates of the system [68]. Here, the observed dynamics of molecular machines consist of large jumps, interleaved with small-scale fluctuations [197]. The large jumps are often taken to indicate a chemical reaction—producing an instantaneous change in the energy potential experienced by the molecular machine—and the full dynamics can then be mapped onto a discrete-state Markov jump model [27, 213]. The observed chemical rates are then used to understand the thermodynamic properties (and functional capabilities) of the mechanochemical machine. However, such an approach by its very nature ignores so-called *hidden* contributions to thermodynamic quantities —such as internal energy flows—arising due to the mechanical dynamics at sub-mesostate resolutions [204]. Such internal energy flows play an important role in understanding the interactions between the components of the molecular machine, and have been used to aid in identifying reaction coordinates in biomolecular dynamics [198].

Independent of such models, much research seeks to understand energy flows into nonequilibrium systems from a time-dependent external perturbation (Sec. 2.6) [106, 107, 109]. These efforts have largely been restricted to deterministic driving protocols—like those typically seen in single-molecule experiments [65, 117, 125, 126, 127, 128]—but have recently been generalized to more closely parallel the *in vivo* dynamics of molecular machines, as in Ch. 6, Ch. 7, and Refs. [180, 190]. However, as of yet, none of these control-theoretical approaches to quantifying energy flows within nanoscale nonequilibrium systems have fully appreciated the essential inter-component feedback present in fully autonomous systems (see Ch. 9) [195, 199].

In this chapter, we develop a thermodynamically complete phenomenological method for quantifying the hidden excess power in strongly coupled nonequilibrium systems. The hidden excess power represents an energy flow communicated between the components of a strongly coupled system, and—unconstrained by the usual form of the second law—can become negative (see Sec. 9.4.2). Such negative excess power is a signature of a fully autonomous Maxwell demon, achieving net flow of heat from the reservoir into a subsystem—where it is transduced into work—as a result of its strong coupling with another driven subsystem.

We find that the hidden excess work per chemical transition can be decomposed into two contributions, the timescale-separated (TSS) excess work (10.17)—the asymptotic contribution to the excess work in the TSS limit—and the nonequilibrium excess work, which is the additional contribution to the excess work when the mechanical states remain out of equilibrium at steady state. The hidden excess power (the excess work per chemical transition, averaged over the coarse-grained dynamics) is typically not easily calculable, as it requires information about hidden states. However, we provide a leading-order approximation of both the TSS excess work (10.20) and nonequilibrium excess work (10.21) which require only minimal information about the hidden-state dynamics. We demonstrate the utility of these approximations in two model systems, representing minimal models of linear and rotary mechanochemical molecular machines.

10.2 Coarse-grained representations of mechanochemical systems

Molecular machines often couple mechanical motion to chemical reactions, for instance in kinesin [214] and F_1-ATP synthase [134]. Thus their dynamics can be described by two coupled coordinates $\boldsymbol{x}$ and $\boldsymbol{\lambda}$, representing the mechanical and chemical degrees of freedom of the system [211, 215]. Each chemical state $\boldsymbol{\lambda}$ induces a particular potential-energy landscape on the mechanical degrees of freedom $E(\boldsymbol{x}|\boldsymbol{\lambda})$ such that, in the absence of any chemical changes, the mechanical state relaxes to a conditional equilibrium distribution

$$\pi(\boldsymbol{x}|\boldsymbol{\lambda}) = e^{-\beta E(\boldsymbol{x}|\boldsymbol{\lambda})+\beta F(\boldsymbol{\lambda})} \ , \tag{10.1}$$

with $\beta \equiv (k_{\mathrm{B}}T)^{-1}$ the inverse temperature of the heat bath and $F(\boldsymbol{\lambda}) \equiv -\ln \int \exp\left[-\beta E(\boldsymbol{x}|\boldsymbol{\lambda})\right] \mathrm{d}x$ the conditional equilibrium free energy given the fixed chemical state $\boldsymbol{\lambda}$.

The dynamics of both the chemical and mechanical states evolve stochastically, with joint transition $(\boldsymbol{\lambda}_i, \boldsymbol{x}') \to (\boldsymbol{\lambda}_j, \boldsymbol{x})$ occurring at rate $R_{ji}^{\boldsymbol{x},\boldsymbol{x}'}$. We further assume that the joint dynamics are bipartite (9.2), so that no simultaneous transitions in both $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ occur ($R_{ji}^{\boldsymbol{x},\boldsymbol{x}'} = 0$ when both $j \neq i$ and $\boldsymbol{x} \neq \boldsymbol{x}'$). Thermodynamically consistent mechanochemical dynamics—relaxing to the correct equilibrium distribution in the absence of any chemical driving—require that the chemical transition rates satisfy generalized detailed balance [205, 206],

$$\frac{R_{ji}^{\boldsymbol{x}}}{R_{ij}^{\boldsymbol{x}}} = e^{-\beta \Delta G_{ji}(\boldsymbol{x})} \ , \tag{10.2}$$

where $R_{ji}^{\boldsymbol{x}} \equiv R_{ji}^{\boldsymbol{x},\boldsymbol{x}}$ indicates the rate of chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ at fixed mechanical state $\boldsymbol{x}$, $\Delta G_{ji}(\boldsymbol{x}) = \Delta\mu_{ji} + \Delta E_{ji}(\boldsymbol{x})$ is the change in free energy during transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, with $\Delta\mu_{ji}$ and $\Delta E_{ji}(\boldsymbol{x})$ the associated respective changes in chemical potential and microstate energy.

We consider transition rates obeying (10.2) of the general form

$$R_{ji}^{\boldsymbol{x}} = \Gamma_{\mathrm{chem}} e^{-\frac{1}{2}\beta[\Delta\mu_{ji}+E(\boldsymbol{x}|\boldsymbol{\lambda}_j)-E(\boldsymbol{x}|\boldsymbol{\lambda}_i)]} \ , \tag{10.3}$$

where Γ_{chem} is a kinetic prefactor (with units of inverse time) that quantifies the bare rate of chemical transitions in the absence of any differences in energy or chemical potential between states.

Specifically, the dynamics of the joint probability $p_j(\boldsymbol{x})$ for system state $(\boldsymbol{\lambda}_j, \boldsymbol{x})$ is governed by the master equation [72]

$$\mathrm{d}_t p_j(\boldsymbol{x}) = \sum_i \int_x R_{ji}^{\boldsymbol{x},\boldsymbol{x}'} p_i(\boldsymbol{x}') \mathrm{d}\boldsymbol{x}' \ , \tag{10.4}$$

where the summation runs over all chemical states $\boldsymbol{\lambda}_i$. Writing the joint probability as $p_j(\boldsymbol{x}') = P_j\, p(\boldsymbol{x}'|\boldsymbol{\lambda}_j)$—where $P_j \equiv \int_{\boldsymbol{x}} p_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ is the marginal distribution of chemical state $\boldsymbol{\lambda}_j$—gives

$$\mathrm{d}_t p_j(\boldsymbol{x}) = \sum_{i,j} P_i \left[\int_{\boldsymbol{x}'} R_{ji}^{\boldsymbol{x},\boldsymbol{x}'} p(\boldsymbol{x}'|\boldsymbol{\lambda}_i)\mathrm{d}\boldsymbol{x}' \right] . \tag{10.5}$$

Integrating over mechanical states $\boldsymbol{x}$ gives the equation of motion for the coarse-grained (chemical) states,

$$\mathrm{d}_t P_j = \sum_{i,j} P_i \int_{\boldsymbol{x}} \left[\int_{\boldsymbol{x}'} R_{ji}^{\boldsymbol{x},\boldsymbol{x}'} p(\boldsymbol{x}'|\boldsymbol{\lambda}_i)\mathrm{d}\boldsymbol{x}' \right] \mathrm{d}\boldsymbol{x} . \tag{10.6}$$

For bipartite dynamics ($R_{ji}^{\boldsymbol{x},\boldsymbol{x}'} = 0$ if $j \neq i$ and $\boldsymbol{x} \neq \boldsymbol{x}'$), the coarse-grained master equation simplifies to

$$\mathrm{d}_t P_j = \sum_{i,j} P_i \int_x R_{ji}^{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{\lambda}_i)\mathrm{d}\boldsymbol{x} \tag{10.7a}$$

$$= \sum_{i,j} V_{ji} P_i , \tag{10.7b}$$

where $R_{ji}^{\boldsymbol{x}} \equiv R_{ji}^{\boldsymbol{x},\boldsymbol{x}}$ is the bipartite transition rate at fixed mechanical state $\boldsymbol{x}$, and $V_{ji} \equiv \int_{\boldsymbol{x}} R_{ji}^{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{\lambda}_i)\mathrm{d}\boldsymbol{x}$ is the coarse-grained rate, the observed transition rate of $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ if the mechanical states are hidden [68].

We view the chemical dynamics as a coarse-grained representation of the mechanochemical system by defining coarse-grained states (*mesostates*) $X_i \equiv \{\boldsymbol{x}|\boldsymbol{\lambda} = \boldsymbol{\lambda}_i\}$ that aggregate all mechanical states at a given chemical coordinate. The coarse-grained rates $V_{ji}(t)$ are generally time-dependent due to the time-dependent conditional microstate distribution $p(\boldsymbol{x}, t|\boldsymbol{\lambda}_i)$, but become time-independent when the mechanical dynamics are at steady state.

In the TSS limit, where at steady state the conditional distribution of mechanical states $p(\boldsymbol{x}|\boldsymbol{\lambda})$ is the conditional equilibrium distribution $\pi(\boldsymbol{x}|\boldsymbol{\lambda})$ (10.1), the dissipation (entropy production) of the joint system is fully determined by the coarse-grained dynamics [68]. However, even in this limit, there are internal flows of energy and information between the mechanical and chemical coordinates. Furthermore, this limit may not hold in general for mechanochemical biomolecular systems—such as molecular machines—which often operate far from thermodynamic equilibrium.

10.3 Hidden excess work in molecular machines

The flow of energy between the chemical and mechanical degrees of freedom can be quantified by the work done on the mechanical system by the chemical dynamics. For a particular chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, the average work done by the chemical coordinate on the mechanical coordinate is $\langle \beta W \rangle_{ji}$, where the average is over the distribution of mechanical states $\boldsymbol{x}$ from which the chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ occurs. The average *excess work* is

$\langle \beta W_{\rm ex} \rangle_{ji} \equiv \langle \beta W \rangle_{ji} - \beta \Delta F_{ji}$, for change $\Delta F_{ji} \equiv F(\boldsymbol{\lambda}_j) - F(\boldsymbol{\lambda}_i)$ in conditional equilibrium free energy upon the chemical transition.

For a system at steady state, the average rate of excess work (the *excess power*) on the mechanical system is expressed in terms of coarse-grained rates as

$$\langle \beta \mathcal{P}_{\rm ex} \rangle_{\Lambda \to X} \equiv \sum_{i,j} V_{ji} P_i \langle \beta W_{\rm ex} \rangle_{ji} \ , \tag{10.8}$$

where angle brackets $\langle \cdots \rangle_{\Lambda \to X}$ indicate the excess power on the mechanical system X averaged over the ensemble of chemical dynamics Λ and system responses.

Here, we provide a method to quantify the hidden contributions to the excess power in a coarse-grained system, by viewing the hidden mechanical states as being driven by the stochastic chemical dynamics. The average work during the chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ is

$$\langle \beta W \rangle_{ji} = \beta \int \Delta E_{ji}(\boldsymbol{x}) \, p_{ji}^{\rm sw}(\boldsymbol{x}) \, {\rm d}\boldsymbol{x} \ , \tag{10.9}$$

where $\Delta E_{ji}(\boldsymbol{x}) \equiv E(\boldsymbol{x}|\boldsymbol{\lambda}_j) - E(\boldsymbol{x}|\boldsymbol{\lambda}_i)$ is the change in energy of $\boldsymbol{x}$ during the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, and $p_{ji}^{\rm sw}(\boldsymbol{x})$ is the distribution of mechanical states $\boldsymbol{x}$ from which the transition occurs (9.14). The switching-state distribution $p_{ji}^{\rm sw}(\boldsymbol{x})$ (9.14) in general depends on the dynamics of both the chemical and mechanical subsystems. When the chemical dynamics are independent of the mechanical state, the switching-state distribution is simply the equilibrium distribution (10.1), and the average TSS work reduces to the *infinite-time work* from Sec.7.3. In this case, the excess work associated with a discrete change in the energy landscape is the relative entropy between the adjacent equilibrium distributions (analogous to (7.9b)). However, for an autonomous system that obeys microscopic reversibility, the chemical transition rates must obey generalized detailed balance (10.2), and the switching-state distribution in general differs from the equilibrium distribution.

Motivated by the approach to discrete control protocols in Ch. 7, we decompose the average excess work for transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ into two components,

$$\langle \beta W_{\rm ex} \rangle_{ji} = \langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} + \langle \beta W_{\rm ex}^{\rm neq} \rangle_{ji} \ , \tag{10.10}$$

where $\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji}$ is the excess work in the timescale-separated (TSS) limit where the mechanical states fully equilibrate between chemical reactions, and $\langle \beta W_{\rm ex}^{\rm neq} \rangle_{ji}$ is the *nonequilibrium excess work* quantifying the additional work beyond the TSS work due to the mechanical degrees of freedom being out of equilibrium.

10.3.1 TSS excess work

For the transition rates in (10.3), the timescale-separated switching-state distribution is the normalized product of the equilibrium distribution and the chemical transition rate:

$$p_{ji}^{\mathrm{sw}}(\boldsymbol{x}) \equiv \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} \pi_i(\boldsymbol{x}) e^{-\frac{1}{2}\beta[E(\boldsymbol{x}|\boldsymbol{\lambda}_j) - E(\boldsymbol{x}|\boldsymbol{\lambda}_i)]} \ , \tag{10.11}$$

where $\mathcal{Z}_{ji}^{\mathrm{sw}}$ is the normalization constant, which contains the prefactors of the rate expression (10.3). Inserting the equilibrium distribution (10.1) gives

$$\begin{aligned} p_{ji}^{\mathrm{sw}}(\boldsymbol{x}) &= \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} e^{-\beta E(\boldsymbol{x}|\boldsymbol{\lambda}_i) - \frac{1}{2}\beta[E(\boldsymbol{x}|\boldsymbol{\lambda}_j) - E(\boldsymbol{x}|\boldsymbol{\lambda}_i)]} && \text{(10.12a)} \\ &= \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} e^{-\frac{1}{2}\beta[E(\boldsymbol{x}|\boldsymbol{\lambda}_i) + E(\boldsymbol{x}|\boldsymbol{\lambda}_j)]} && \text{(10.12b)} \\ &= \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} \sqrt{\pi_i \pi_j} \ , && \text{(10.12c)} \end{aligned}$$

the normalized geometric mean of the equilibrium distributions at $\boldsymbol{\lambda}_i$ and $\boldsymbol{\lambda}_j$. Thus the normalization factor is

$$\mathcal{Z}_{ji}^{\mathrm{sw}} = \int \sqrt{\pi_i \pi_j} \, \mathrm{d}\boldsymbol{x} \ , \tag{10.13}$$

which in information theory is the *Bhattacharyya coefficient*, a measure of the difference between two probability distributions [216].

Substituting in (10.9) the energies in terms of the corresponding equilibrium probabilities,

$$\beta E(\boldsymbol{x}|\boldsymbol{\lambda}_i) = -\ln \pi_i(\boldsymbol{x}) + \beta F(\boldsymbol{\lambda}_i) \ , \tag{10.14}$$

and the equilibrium switching-state distribution (10.12c), the TSS work becomes

$$\langle \beta W^{\mathrm{TSS}} \rangle_{ji} = \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} \int \sqrt{\pi_i \pi_j} \, \ln \frac{\pi_i}{\pi_j} \, \mathrm{d}\boldsymbol{x} + \beta \Delta F_{ji} \ , \tag{10.15}$$

where $\Delta F_{ji} \equiv F(\boldsymbol{\lambda}_j) - F(\boldsymbol{\lambda}_i)$ is the equilibrium free energy difference between chemical states $\boldsymbol{\lambda}_i$ and $\boldsymbol{\lambda}_j$. Thus the *excess work* is

$$\begin{aligned} \langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{ji} &\equiv \langle \beta W^{\mathrm{TSS}} \rangle ji - \beta \Delta F_{ji} && \text{(10.16a)} \\ &= \frac{1}{\mathcal{Z}_{ji}^{\mathrm{sw}}} \int \sqrt{\pi_i \pi_j} \, \ln \frac{\pi_i}{\pi_j} \, \mathrm{d}\boldsymbol{x} \ . && \text{(10.16b)} \end{aligned}$$

The transition-specific TSS excess work also equals a difference between relative entropies:

$$\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} = \frac{1}{\mathcal{Z}_{ji}^{\rm sw}} \int \sqrt{\pi_i \pi_j} \, \ln \frac{\pi_i}{\pi_j} \, \mathrm{d}\boldsymbol{x} \tag{10.17a}$$

$$= \frac{1}{\mathcal{Z}_{ji}^{\rm sw}} \int \left(\ln \frac{\sqrt{\pi_i \pi_j}}{\pi_j} - \ln \frac{\sqrt{\pi_i \pi_j}}{\pi_i} \right) \sqrt{\pi_i \pi_j} \, \mathrm{d}\boldsymbol{x} \tag{10.17b}$$

$$= D\left(\sqrt{\pi_i \pi_j} \,||\, \pi_j\right) - D\left(\sqrt{\pi_i \pi_j} \,||\, \pi_i\right) \ . \tag{10.17c}$$

Intuitively, the TSS excess work quantifies the increase in additional free energy of the switching-state distribution $p_{ji}^{\rm sw}(\boldsymbol{x})$ through the chemical reaction $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ [217]. (Eq. (10.17c) was also derived for master-equation dynamics for a general switching-state distribution in Ch. 9.)

The TSS excess work (10.17c) is antisymmetric under index exchange, $\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} = -\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ij}$, and thus the TSS excess power (10.8) (obtained by summing the transition-specific excess work over coarse-grained dynamics) is

$$\langle \beta \mathcal{P}_{\rm ex}^{\rm TSS} \rangle_{\Lambda \to X} = \sum_{i<j} \left(P_i V_{ji} - P_j V_{ij} \right) \langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} \ . \tag{10.18}$$

Thus, even for nonzero $\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji}$, the associated hidden excess power vanishes when there is no net probability flux in the chemical coordinate. This occurs, for instance, when there is no chemical driving ($\Delta\mu = 0$), and is a direct result of the autonomous nature of the system, exemplifying the essential importance of requiring the rates (10.3) to satisfy local detailed balance (10.2).

The TSS excess work further simplifies in the small-$\Delta\boldsymbol{\lambda}$ limit, where Taylor expanding (10.16b) produces (see App. G.1)

$$\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} \approx \frac{1}{4!} \mathcal{S}_{m\ell k}(\boldsymbol{\lambda}_i) \Delta\lambda_{ji}^m \Delta\lambda_{ji}^\ell \Delta\lambda_{ji}^k \ , \tag{10.19}$$

where λ_{ji}^k is the change in the kth component of the chemical state vector $\boldsymbol{\lambda}$ during the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, and the rank-3 tensor

$$\mathcal{S}_{m\ell k}(\boldsymbol{\lambda}) = \langle \delta f_m \, \delta f_\ell \, \delta f_k \rangle_{\boldsymbol{\lambda}_i} \ , \tag{10.20}$$

is the third centered moment of conjugate forces $f_k \equiv -\partial_{\lambda^k} E(\boldsymbol{x}|\boldsymbol{\lambda})$. Here, we use an Einstein summation notation, where repeated indices are implicitly summed over. In contrast, the analogous infinite-time excess work in [77] is second order in $\Delta\boldsymbol{\lambda}$ and uses the force variance $\langle \delta f_m \delta f_k \rangle_{\boldsymbol{\lambda}_i}$ instead of the third centered moment. Here, the feedback of the mechanical system on the chemical dynamics results in cancellation of the second-order terms.

10.3.2 Nonequilibrium excess work

In the small-$\Delta\boldsymbol{\lambda}$ limit, linear-response theory provides a simple approximation for the nonequilibrium excess work $\langle \beta W_{\mathrm{ex}}^{\mathrm{neq}} \rangle_{ji}$. In particular, if upon the chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, the corresponding change in the mechanical energy landscape $E(\boldsymbol{x}|\boldsymbol{\lambda}_i) \to E(\boldsymbol{x}|\boldsymbol{\lambda}_j)$ is small, $E(\boldsymbol{x}|\boldsymbol{\lambda}_j)$ is well approximated by a first-order Taylor expansion about $\boldsymbol{\lambda}_i$. We additionally assume a moderate timescale separation between the mechanical and chemical dynamics, such that the relaxation time of the conjugate forces f_k is significantly shorter than the chemical-state dwell time. This constraint is weaker than the timescale-separation limit typically considered in coarse-graining, where the hidden states are assumed to relax infinitely faster than the chemical dynamics and thus equilibrate fully [68].

Under these assumptions, the leading-order contribution to the nonequilibrium excess work during chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ is (see App. G.2)

$$\langle \beta W_{\mathrm{ex}}^{\mathrm{neq}} \rangle_{ji} \approx \beta \Gamma_{ji} \zeta_{kk'}(\boldsymbol{\lambda}_i) \Delta\lambda_{ji}^{k} \frac{\sum_s P_s V_{is} \Delta\lambda_{is}^{k'}}{V_{i*}} \ . \tag{10.21}$$

$\Gamma_{ji} \equiv \Gamma_{\mathrm{chem}} \exp\left(\frac{1}{2}\Delta\mu_{ji}\right)$, $V_{i*} \equiv \sum_s P_s V_{is}$ is the total rate of chemical transitions entering $\boldsymbol{\lambda}_i$, and the summation index s runs over all chemical states. $\zeta_{kk'}(\boldsymbol{\lambda}) \equiv \beta \int_0^\infty \langle \delta f_k(0) \delta f_{k'}(t) \rangle_{\boldsymbol{\lambda}} \, \mathrm{d}t$ is the *generalized friction tensor* [109] that quantifies the dissipation due to nonequilibrium driving by a deterministic control protocol. Equation (10.21) quantifies the nonequilibrium excess work using only information about the coarse-grained dynamics and the friction tensor $\zeta_{kk'}(\boldsymbol{\lambda})$, which is determined from the hidden states' equilibrium fluctuations.

At steady state, the excess power by the chemical dynamics on the mechanical system, averaged over the coarse-grained transitions, is

$$\langle \beta \mathcal{P}_{\mathrm{ex}} \rangle_{\Lambda \to X} = \sum_{i,j} V_{ji} P_j \left[\langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{ji} + \langle \beta W_{\mathrm{ex}}^{\mathrm{neq}} \rangle_{ji} \right] \ . \tag{10.22}$$

The near-equilibrium expressions for the TSS (10.19) and nonequilibrium excess work (10.21) in terms of conditional equilibrium averages permits estimation of the excess power in experimental investigation of autonomous mechanochemical molecular machines: the low-resolution observations of hidden-state fluctuations it requires are more tractable than inferring the full mechanochemical dynamics of the motor [40, 218].

10.4 Model systems

We illustrate our theoretical predictions by investigating the hidden excess power in two model systems, representing linear-transport and rotary molecular motors. In both cases, the chemical degrees of freedom evolve on a discrete, one-dimensional lattice, with rates given by (10.3), while the mechanical coordinate x (or θ in the rotary model) diffuses in

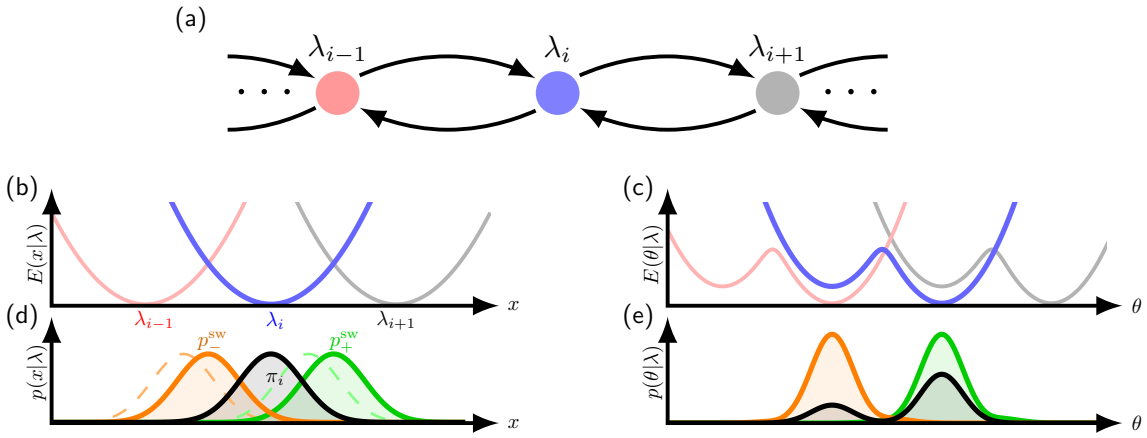

Figure 10.1: **Schematic of mechanochemical motor.** (a) Chemical reaction network, an infinite one-dimensional lattice of chemical states. (b-c) Conditional energy landscapes for (b) linear-transport $E(x|\lambda)$ and (c) rotary $E(\theta|\lambda)$ motors. (d-e) Equilibrium distribution (black) and forward/reverse equilibrium switching-state distributions (solid green/orange) for λ_i (blue state in (a) and blue potential in (b,c)) for both the (d) linear-transport and (e) rotary motors. Dashed curves in (d) indicate the nonequilibrium switching-state distributions for a system in which the chemical coordinate has a net flux towards higher index λ values (to the right in (a)), thus lagging behind their corresponding equilibrium switching-state distributions.

a one-dimensional energy landscape $E(x|\lambda)$ (or $E(\theta|\lambda)$) determined by the instantaneous chemical state λ. Figure 10.1 shows a schematic of both model systems.

In the linear-transport motor, the TSS excess work (10.17) is zero, thus serving as a model to isolate the nonequilibrium excess work (10.21). Conversely, for the rotary motor we analyze its excess work in the TSS limit, where the nonequilibrium excess work is zero, thereby isolating the effects of the TSS excess work (10.17).

10.4.1 Linear-transport motor

The linear-transport motor has a one-dimensional mechanical degree of freedom subject to a harmonic potential $E_{\mathrm{harm}}(x|\lambda_i) = \frac{1}{2}k_{\mathrm{trap}}(x-\lambda_i)^2$, with the instantaneous chemical state λ_i determining the potential minimum. Figure 10.1(a,b,d) shows a schematic.

The chemical coordinate evolves on an infinite one-dimensional lattice of chemical states, with only nearest-neighbor transitions. Because the adjacent equilibrium distributions are only distinguished by their mean, we observe the system in a comoving frame by changing coordinates to the relative position $x - \lambda_i$. This reference frame has a steady-state distribution for time-independent transition rates (10.3) and uniform fixed $\Delta\mu_{i+1,i} = -\Delta\mu < 0$ biasing the machine to (on average) move forward. Here, the generalized friction is uniform and equals the viscous friction, $\zeta = \gamma$, and the steady-state coarse-grained forward and reverse rates $V_\pm$ and probabilities P_i are uniform. Equation (10.21) approximates the

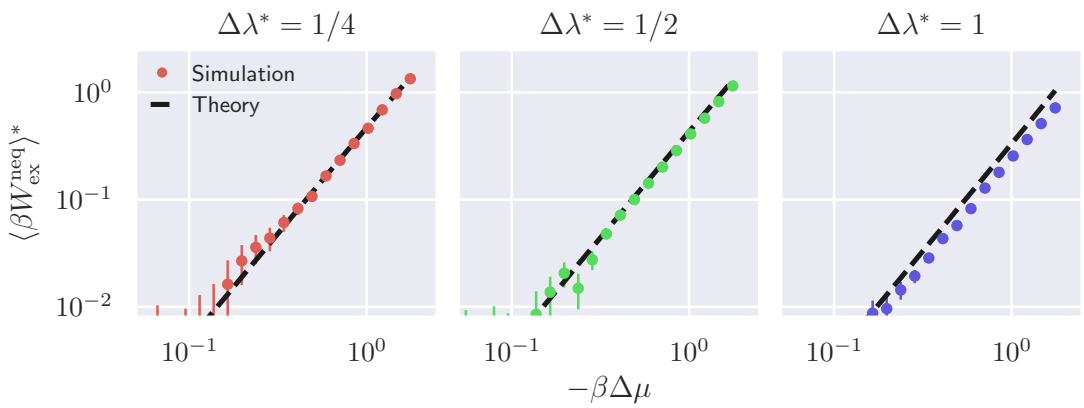

Figure 10.2: **Nonequilibrium excess work for linear-transport motor.** Nonequilibrium excess work $\langle \beta W_{\rm ex}^{\rm neq} \rangle^* \equiv \langle \beta W_{\rm ex}^{\rm neq} \rangle / [\beta \Gamma_{\rm chem} \zeta \Delta\lambda^2]$, as a function of forward chemical-potential bias $-\beta\Delta\mu$, for numerical simulation (dots) or approximate theory (10.24) (black dashed line). Different panels show different step sizes, non-dimensionalized by the equilibrium standard deviation $\sigma_x = 1/\sqrt{\beta k_{\rm trap}}$ of the mechanical coordinate at fixed λ : $\Delta\lambda^* \equiv \Delta\lambda/\sigma_x$.

steady-state nonequilibrium excess work, and in this case can be greatly simplified. The average excess work for a forward (+) or reverse (−) step is

$$\langle \beta W_{\rm ex}^{\rm neq} \rangle_{\pm} \approx \pm \beta\zeta\Gamma_{\pm}\Delta\lambda^2 \frac{V_+ - V_-}{V_+ + V_-} \ . \tag{10.23}$$

Here, $\Gamma_{\pm} \equiv \Gamma \exp(\pm\frac{1}{2}\beta\Delta\mu)$, for chemical potential difference $\beta\Delta\mu$ for a forward chemical step. The $\pm$ prefactor in (10.23) reflects that, on average, forward steps require positive work input while reverse steps require negative work input. Thus the average work done for any chemical transition (forward or reverse) is the mean of $\langle \beta W_{\rm ex}^{\rm neq} \rangle_+$ and $\langle \beta W_{\rm ex}^{\rm neq} \rangle_-$, each term weighted by the corresponding $\pm$ jump probabilities:

$$\langle \beta W_{\rm ex}^{\rm neq} \rangle_{\Delta\lambda} \approx \frac{V_+ \langle \beta W_{\rm ex}^{\rm neq} \rangle_+}{V_+ + V_-} + \frac{V_- \langle \beta W_{\rm ex}^{\rm neq} \rangle_-}{V_+ + V_-} \tag{10.24a}$$

$$= \beta\zeta\Delta\lambda^2 \frac{V_+ - V_-}{(V_+ + V_-)^2} \left(V_+\Gamma_+ - V_-\Gamma_- \right) \ . \tag{10.24b}$$

Here, the nonequilibrium excess work depends on the chemical driving $\Delta\mu$ through both the coarse-grained chemical rates $V_{\pm} \propto \exp\left(\pm\frac{\beta}{2}\Delta\mu\right)$ (following from their dependence on microscopic transition rates R^x_{ji} and (10.3)) and $\Gamma_{\pm} \propto \left(\pm\frac{\beta}{2}\Delta\mu\right)$. Figure 10.2 compares the theoretical predictions (10.24) with numerical results, showing good agreement in the small-$\Delta\lambda$ limit.

10.4.2 Rotary motor

Next, we consider the Kawaguchi-Sasa-Sagawa (KSS) model of F_1-ATPase [134]. Here, the rotational angle θ of F_1-ATPase's crankshaft evolves in the potential of mean force (PMF) discussed in Sec. 2.7.3

$$\beta E_{\mathrm{KSS}}(\theta|\lambda_i) = \tfrac{1}{2}\beta k_{\mathrm{c}}\left(\theta - \lambda_i\right)^2 - \ln\left[e^{-\beta k_{\mathrm{c}}\phi(\theta-\lambda_i)} + e^{\beta\Delta E_{\mathrm{S}} + \frac{1}{2}\beta k_{\mathrm{c}}\phi^2}\right] , \tag{10.25}$$

arising from fast switching between two harmonic potentials, each with stiffness k_{c}, with angular offset ϕ and energetic offset ΔE_{S}. Figure 10.1c shows a schematic of the potential (10.25). Unlike the (purely harmonic) linear-transport model (Sec. 10.4.1), for $\Delta E_{\mathrm{S}} \neq 0$ and $\phi \neq 0$ this potential is asymmetric in θ, generally producing non-zero TSS excess work (10.16b). To isolate the effects of the nonzero TSS excess work, we consider this model in the TSS limit, where the mechanical states come to conditional equilibrium between each chemical transition, and thus there is no contribution from the nonequilibrium excess work (10.21).

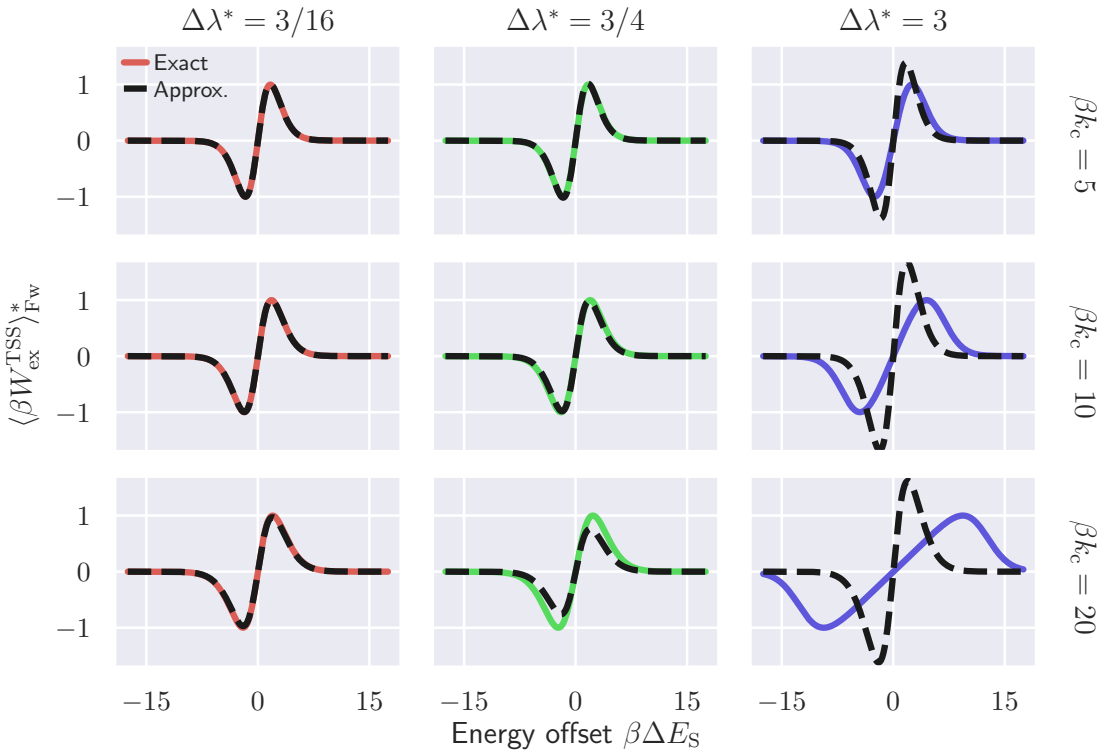

Figure 10.3: **TSS excess work in rotary motor.** TSS work $\langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle^*_{\mathrm{Fw}} \equiv \langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{\mathrm{Fw}} / \max_{\Delta E_{\mathrm{S}}} \langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{\mathrm{Fw}}$ per forward step, as a function of energy offset ΔE_{S}, for several chemical step sizes $\Delta\lambda^* \equiv \Delta\lambda/\phi$ (columns), nondimensionalized by the angular offset between minima, several and spring constants βk_{c} (rows). Colored solid curves: exact; dashed curves: small-$\Delta\lambda$ approximation (10.19).

Figure 10.3 compares the exact TSS excess work $\langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{\mathrm{Fw}}$ per forward step (10.17) with its leading-order approximation (10.19), as a function of the energy offset ΔE_{S}. For small (but nonzero) energy offset ΔE_{S}, the switching-state distribution (10.12) for forward steps is heavily concentrated in the right minimum of the potential (Fig. 10.1c), where for positive (negative) ΔE_{S} the changes in energy of the mechanical state $\boldsymbol{x}$ upon the chemical transition are positive (negative). For large-magnitude offsets $|\Delta E_{\mathrm{S}}|$, the potential becomes effectively harmonic, and thus the TSS work (10.17) vanishes. For smaller $\Delta\lambda/\phi$ and lower βk_{c}, the approximate form of the TSS excess work (10.19) agrees well with the exact calculation (10.17). For large step sizes and stronger spring constants, the approximation in (10.19) begins to break down.

At the level of the (driven) mechanical subsystem, negative TSS excess work indicates net heat flow from the thermal reservoir into the mechanical subsystem and onward as work extracted by the chemical dynamics. The strong coupling between mechanical and chemical subsystems ensures thermodynamic consistency by a compensating heat flow from the chemical coordinate to the thermal reservoir, rescuing what would otherwise be a violation of the second law [195, 199]. In such circumstances, the chemical dynamics operate as a fully autonomous Maxwell demon, effectively using the information gleaned through strong coupling to the mechanical system to transduce heat from the bath into extracted work.

10.5 Discussion

In this chapter, we present a phenomenological formalism to estimate the hidden excess power internal to coarse-grained autonomous nonequilibrium systems, which only requires the coarse-grained chemical dynamics and minimal information about the hidden mechanical dynamics. This theoretical framework provides a means to estimate the hidden energy flows within molecular machines without explicitly modeling their microscopic dynamics.

We identify two distinct contributions to the hidden excess power: the TSS excess work (10.17)—which persists in the TSS limit, when between chemical reactions the hidden mechanical states fully relax to a conditional equilibrium—and the nonequilibrium excess work (10.21), which is the additional energetic cost due to the mechanical states being out of equilibrium.

The exact TSS excess work $\langle \beta W_{\mathrm{ex}}^{\mathrm{TSS}} \rangle_{ji}$ for a particular chemical transition is the difference between two relative entropies (10.17), and can be interpreted as quantifying the concomitant change of the additional free energy in the mechanical system (see Ch. 9). Unlike the analogous infinite-time excess work in [77], the TSS excess work can be either positive or negative. This is consistent with findings in Ch. 9, where we showed that—due to the strong coupling between the mechanical and chemical coordinates—the excess power (or rate of excess work) $\langle \beta \mathcal{P}_{\mathrm{ex}} \rangle_{\Lambda \to X}$ is not, in itself, constrained by the second law, and thus can become negative.

We also identify a phenomenological expression that approximates (to leading order in $\Delta\boldsymbol{\lambda}$) the TSS work of the chemical dynamics on the mechanical system by the third centered moment of the conditional equilibrium distribution of conjugate forces (10.20). This is in contrast to a similar expansion of the infinite-time excess work in Ch. 7 that is second-order in $\Delta\boldsymbol{\lambda}$ and uses the force variance in place of the third moment in (10.20). Here, feedback from the mechanical state to the chemical dynamics leads to exact cancellation of terms second-order in $\Delta\boldsymbol{\lambda}$. Additionally, in the moderate timescale-separated limit (Sec. 10.3.2), the nonequilibrium excess work is determined by the generalized friction tensor (10.21), which can be inferred through observations of equilibrium force fluctuations.

Our theoretical framework can be used as a tool in experimental studies of molecular machines to better understand excess power in autonomous systems. In particular, recent experiments on kinesin [218] and F_1 ATPase [40] have used their mechanical fluctuations to infer hidden entropy production; these measurements could also be used to calculate both the TSS excess work and nonequilibrium excess work.

Furthermore, the identification of negative steady-state excess power has an interesting physical interpretation, providing a signature of Maxwell-demon behavior in a mechanochemical machine [199]. Specifically, negative excess power means that, on average, there is a net heat flow from the reservoir into the subsystem being driven. We showed an example of this with the TSS excess work in the rotary model system (Sec. 10.4.2), and analogous behavior was recently found in a similar bistable potential in Sec. 9.4.2. Our framework provides a tractable method to measure this quantity, and can be used to identify such thermodynamic properties in biological molecular machines.

Chapter 11

Conclusions and outlook

Over the past two and a half decades, the emergence of stochastic thermodynamics and fluctuation theorems has bolstered an increasingly deep understanding of the physics of fluctuating systems. However, quantifying the flow of energy, information, and entropy within and through fluctuating systems remains a central challenge in our understanding of biology at the microscopic scale. Molecular machines represent a particular class of microscopic biological system which have garnered significant attention within the biophysics community in recent years; however, our understanding of their physical limitations is still developing. Recent advances such as, for instance, the thermodynamic uncertainty relation introduced by Barato and Seifert in 2015 [29, 30] and its subsequent generalizations [31, 32, 33] have shed light on some of the fundamental limitations that are present in such systems. However, much work is still required.

In this thesis we have presented a series of results furthering our understanding of the physics of biological systems at the microscopic scale. Our approach primarily builds upon a view of such systems through the lens of control theory (Sec. 2.6), with systems driven through active manipulation of control parameters. Here, we seek to understand the response of a stochastic system to a time-dependent perturbation, and how that response relates to its dissipation. Such a framework was the basis for many significant developments, such as the original derivation of the Jarzynski equality [6], the Crooks fluctuation theorem [7], and the geometric interpretations of thermodynamic systems [120, 121, 122, 124]. Whether it be in the classical sense of Weinhold [120] or Ruppeiner [121], or the more modern formulation of thermodynamic geometry in terms of the Fisher information [124] or the generalized friction [109], geometric descriptions of thermodynamic systems have clarified and structured many of the problems that we have focused on throughout this thesis. In particular, the geometric view of thermodynamic processes permits a simple interpretation of minimum-dissipation driving protocols as the geodesic curves—in the space of control parameters—connecting initial and final states [109, 123, 124].

While such theoretical descriptions are useful in framing the problem of quantifying dissipation in out-of-equilibrium systems, those mentioned make near-equilibrium approxi-

mations, raising questions about their validity for *in vivo* molecular machines, which often operate far from thermodynamic equilibrium. In Part I of this thesis, we explored the utility of the generalized-friction framework (Sec. 2.6.2) for designing reduced-dissipation control protocols in a model biophysical system: a DNA hairpin (Fig. 3.1). In Chapter 3 we discussed the estimation of the generalized friction as a function of trap separation for a DNA hairpin held between two polystyrene beads that are both held in an optical trap. Here, we measured equilibrium force fluctuations of the DNA hairpin construct (Fig. 3.2) at various trap separations to obtain an estimate of the generalized friction coefficient (Fig. 3.3) $\zeta(\lambda)$ (2.65) which showed a prominent peak at the separation $\lambda_{1/2}$ where at equilibrium it is equally probable for the hairpin to be in the folded or unfolded state. Using our estimate of the generalized friction, we made use of a simple proportionality (2.72) predicted by theory to design protocols that reduce the excess work required to unfold or refold the hairpin, relative to a naive, constant-velocity protocol of the same duration (Fig. 3.4). The designed protocols proceed comparatively slowly where the friction is large, and thus, intuitively, allow the hairpin sufficient time near $\lambda_{1/2}$ for thermal fluctuations to naturally unfold or refold the hairpin structure.

In Chapter 4, we experimentally implemented the reduced-dissipation protocols designed in Chapter 3. First, we discussed the identification of a unique unfolding/refolding force for each individual force-separation curve (Fig. 4.2), and argued that—given the unfolding/refolding force is a proxy for the distance from equilibrium of the hairpin—that designed protocols should exhibit lower unfolding and higher refolding forces (on average) than their naive counterparts. These theoretical predictions are borne out in Fig. 4.3b, showing that the difference in unfolding/refolding force between designed and naive protocols increases as the protocol duration decreases (i.e., as the average protocol velocity increases), and the hairpins are pushed further out of equilibrium. This analysis is complemented by our calculation of excess power as a function of separation (Fig. 4.5), showing—in agreement with theoretical predictions—that naive protocols require comparatively high excess power near $\lambda_{1/2}$ where the friction peaks, whereas the designed protocols effectively flatten the accumulation of excess power, evenly accumulating it over all separations. We also compared the cycle work (or hysteresis) of unfolding/refolding protocol pairs of a particular class (designed or naive) and found that the designed protocols systematically reduce the excess work required to complete an unfolding-refolding cycle of the DNA hairpin across protocol durations varying by a factor of ~ 30 (Fig. 4.6). Finally, we looked at a theoretical prediction of the ratio of excess works along a naive protocol versus a designed protocol and found that in the long-duration limit the ratio agrees (within experimental error) with the theoretical prediction (Fig. 4.7).

In Chapter 5, we discussed how our results generalize to different hairpin sequences and experimental conditions. In particular, we showed experimental results obtained for an alternative hairpin sequence, referred to as the 'slow-relaxing hairpin', which showed

significantly slower relaxation near $\lambda_{1/2}$ due to base pair modifications in the hairpin stem (Fig. 5.1). By performing analysis of the equilibrium fluctuations—in a similar manner to Chapter 3—we find that the generalized friction is increased at $\lambda_{1/2}$ by a factor of ~ 100 relative to the 'fast-relaxing' hairpin sequence considered in Chapter 3 (Fig. 5.1c). Ultimately, this disparity in the friction magnitude is due to an increased relaxation time, as the force variance between the two hairpins is largely unchanged (Fig. 5.1a). However, the resulting designed protocols appear qualitatively unchanged between the two hairpins, both exhibiting comparatively slow velocities in the regime near $\lambda_{1/2}$ and comparatively large velocities for $\lambda \gg \lambda_{1/2}$ or $\lambda \ll \lambda_{1/2}$ (Fig. 5.1d,e). Our analysis of the cycle work in the slow-relaxing hairpin found similar results to those in Chapter 4, with designed protocols systematically requiring less excess work than their naive counterparts. We also varied the buffer conditions during equilibrium sampling by adding a divalent cation (Mg^{2+}) known to facilitate the formation of alternative hairpin secondary structures. Here, we found that our estimations of the generalized friction remained unchanged, indicating that the designed protocols are insensitive to such variations in environmental conditions.

In Part II of this thesis, we investigated several extensions of the linear-response framework outlined in Section 2.6.2. While the results in Part I justify the utility of linear-response frameworks within the context of biomolecules, the control protocols implemented by molecular machines differ significantly from the protocols implemented in experiments. In particular, the control protocols implemented by molecular machines are often produced by a series of chemical reactions, and thus the protocols are both discrete in space (due to the difference in timescales associated with chemical reaction and mechanical response), and stochastic (due to the stochastic nature of chemical reaction dynamics). However, in its original presentation, the generalized-friction framework (Sec. 2.6.2) assumes control protocols which are both deterministic and continuous. In Chapter 6, we investigated the leading-order corrections to this theory when the control protocols are stochastic. Ultimately, this leads to a competition between two different contributions to the average excess work that is minimized for a finite average protocol duration. This indicates that such systems have a thermodynamic justification for rapid operation.

In Chapter 7, we investigated the linear-response approximation for discrete control protocols. Here, we found that the excess work could be decomposed into two contributions, the infinite-time work—a contribution to the excess work which occurs even in the infinite-time limit, solely due to the discrete nature of the protocols—and the nonequilibrium excess work, which is the leading-order contribution to the excess work when the protocol takes place during a finite (but long) duration. For a harmonic trap driven by a discrete protocol, we found an exact solution (7.52) for the excess work, which agrees quantitatively with the approximate form (7.53)—obtained through the linear-response formalism—in the long-duration limit (Fig. 7.1). We also used the discrete-protocol formalism to numerically obtain minimum-work protocols in a periodic potential, finding that in general such protocols take

large jumps to avoid regions where the force variance is large, and spend longer times in regions where the force autocovariance is slowly decaying (Fig. 7.3). However, as the number of discrete steps in the protocol increases, the minimal-work protocols found using the discrete excess work become equivalent to minimal-work protocols obtained for a continuous protocol using the generalized friction (Fig. 7.5).

In Chapter 8, we combined the ideas presented in Chapters 6 and 7 to derive a formalism capturing the energetic costs associated with discrete, stochastic control-protocol ensembles. By adding an effective cost of control based on the asymmetry of control-parameter dynamics, we connected the resulting theoretical framework with recent work by Bryant and Machta in [180, 190]. Here, we found different lower bounds on the total excess work under differing constraints on the relative timescales of system and control-parameter dynamics. In the time-scale separated limit (Sec. 8.4.1), where the system comes to equilibrium between each discrete control-parameter jump, the lower bound on excess work from Ref [180] applies, while when the system is pushed out of equilibrium, a lower bound similar to the one discussed in Ch. 6 is applicable (Sec. 8.4.2).

However, in all cases considered in this part of the thesis (Chapters 6–8) we consistently assumed that the control-parameter dynamics received no feedback from the system. In the following part (Part III), we discuss why this assumption breaks certain constraints on the microscopic dynamics present in thermodynamically consistent molecular machines. Thus, the applicability of such dissipation bounds is seemingly limited in the systems for which they were originally motivated. However, the mathematical frameworks within which they are derived are consistent, and may serve as important frameworks to understand the trade-offs in dissipative costs associated with different types of systems.

Finally, in Part III of the thesis, we looked at quantifying the flows of energy and entropy within strongly coupled thermodynamic systems. Here, we explored the view of molecular machines as being driven by a time-dependent control parameter within the broader context of stochastic thermodynamics. In Chapter 9 we explored the relationship between excess power and entropy production in strongly coupled systems. This was motivated by the equality between excess work and entropy production—subject to the appropriate boundary conditions—in systems driven by an external control parameter. Here, the equality between excess work and entropy production is encoded in the equivalence of the Crooks fluctuation theorem [7] and the entropy production fluctuation theorem [12] for such systems. We found that in multi-component strongly coupled systems, the excess work done by one component on another does not have any direct relationship to an entropy production. While this excess power can be equated to an entropy production rate in the reservoir, it is not, in itself, constrained by the usual form of the second law, and thus can become negative. We introduced the transduced additional free energy rate (TAFER) as an alternative measure of dissipation that (similar to the excess power) is calculated by averaging over upstream dynamics, and is equal to an entropy production rate. This finding is particu-

larly important for studying the dissipation in molecular machines, as it reveals that the inter-system coupling required by the detailed balance condition is an essential factor in understanding the physical constraints on a particular machine.

In Chapter 10, we focused on quantifying the flow of energy between chemical and mechanical degrees of freedom in a mechanochemical molecular machine. Here, we used linear-response theory to derive a phenomenological expression for the excess power done by the chemical reaction dynamics on the mechanical coordinate in a thermodynamically complete framework. Motivated in large part by the decomposition of excess work into asymptotic and nonequilibrium contributions in Ch. 7, we showed that the excess work can be written as a sum of the timescale-separated excess work—which persists even in the timescale-separated limit—and the nonequilibrium excess work, which quantifies the additional energetic flows when the mechanical system remains out of thermodynamic equilibrium. In each case, we derived a phenomenological expression to approximate the total excess work in a moderate time-scale separated limit, which is analogous to the near-equilibrium approximation made in Ch. 7. Here—and as rigorously outlined in Ch. 9—we found that the excess work can become negative (Sec.10.4.2), which indicates the presence of an implicit Maxwell demon in the molecular machine. A nonzero information flow from the mechanical system back into the chemical dynamics—due to the inter-system feedback required by local detailed balance—results in a net negative excess work (due to a net heat flow into the mechanical system from the thermal reservoir) done by the chemical dynamics when driving the mechanical system. Ultimately, our derivations provide a simple phenomenological expression to quantify the hidden excess power, which can be calculated from equilibrium fluctuations alone, and thus can be used immediately in single-molecule experiments—such as those outlined in [40] or [218]—to better understand the internal energy flows of biological molecular machines.

11.1 Outlook

Looking forward, the results obtained in this thesis suggest a number of interesting future directions of study. The results in Part I show that—although exceedingly simple in its construction—the generalized-friction framework can exhibit significant utility in the design of nonequilibrium protocols that reduce excess work. Thus, future experiments aimed at measuring the generalized friction coefficient in sophisticated molecular machines could be used to better understand their *in vivo* operation. Electrorotation studies of the F_1 component of ATP synthase have already been used to understand its energetic structure [135] and operational efficiency [40, 63]. These experiments could measure equilibrium torque fluctuations in the motor, as a function of angular orientation, and thus estimate the generalized friction as a function of rotation angle. Seen through the lens of control theory, the *in vivo* operation of F_oF_1 ATP synthase consists of the F_o component—driven by the passage

of protons across the mitochondrial membrane—pushing on the F_1 component. Thus the generalized friction coefficient of the F_1 component approximates the dissipative costs faced by F_o during its natural operation. We conjectured in Chapter 5 that, given energetic efficiency has provided an evolutionary pressure for biological molecular machines [129], these minimum-dissipation protocols may bear resemblance to, for instance, the *in vivo* protocols imparted by F_o on F_1.

The incorporation of measurement-based feedback is another potentially interesting direction in which to use our approach to the unfolding and refolding of DNA hairpins. This could lead to further reductions in the amount of total dissipation required to unfold/refold a DNA hairpin, possibly even leading to work extraction. For instance, recent work by Ribezzi-Crivellari and Ritort [219, 220] showed how one can use continuous feedback to implement a Maxwell demon in a DNA hairpin. Specifically, by observing whether the hairpin has spontaneously unfolded, and applying a subsequent control protocol contingent on the readout of such observations, one can rectify fluctuations from the thermal reservoir. This is one of a growing body of experimental realizations of Maxwell demons [16, 221, 222, 223] cast within the emerging field of information thermodynamics [224]. An interesting future direction would be to observe how the dissipation involved in unfolding or refolding a DNA hairpin could be reduced through the implementation of feedback in the unfolding or refolding protocols. For instance, along a single force-separation curve for an unfolding protocol—such as those in the left-hand column of Fig. 4.1—instead of implementing a fixed control protocol, independent of the response of the system, it would be interesting to instead adaptively use observations of whether or not the hairpin has unfolded to consequently speed up or slow down the trap velocity, depending on the measurement outcome. We expect that, in such experiments, implementing some form of designed protocols would be beneficial, as they will tend to allocate a significant fraction of their total duration in regions where the hairpin is likely to experience hopping events, and thus will increase the probability that, along a given force-separation curve, a favourable fluctuation of the hairpin will be observed.

While the extensions to the generalized-friction framework in Part II represent what we came to refer to as thermodynamically incomplete systems, they may yet prove useful in the study of active matter. In our analysis of the disparity between thermodynamically complete and incomplete systems in Ch. 9, we found that driving dynamics which are independent of the state of the system violate local detailed balance, and lead to a finite entropy production rate under equilibrium conditions. Ultimately, this is the reason that we found a set of lower bounds on dissipation in Ch. 8, however, we also saw (in Ch. 9) qualitatively similar conclusions for any value of the feedback parameter α which does not correspond to detailed balance. By one definition, active matter is the study of statistical properties of agents which break detailed balance locally, due to their own internal energy-generation mechanisms [225].

Thus, the lower bounds and general results found in Part II may serve an important role, for instance, in how active materials interact with and manipulate their environments.

We showed in Part III how to quantify energy flows communicated between the components of a strongly coupled mechanochemical system, and found that it was possible to approximate this energy—in a moderate timescale-separated limit—with simple phenomenological expressions for the TSS and nonequilibrium excess powers. Entropy production is an important performance measure in the physics of molecular machines, and in Ch. 9 we showed that, in terms of upstream dynamics, the entropy production rate is equal (at steady state) to the transduced additional free energy rate (TAFER), and not the excess power. Thus, a natural followup question is to ask how one could use similar methodologies to approximate the TAFER directly. This could be done either by supplementing the framework built in Ch. 10 with an approximation for the information rate $\dot{I}^Y$, or by directly approximating the TAFER as a rate of change of additional nonequilibrium free energy, as in (9.25).

Additionally, a detailed numerical study of the TAFER and hidden excess power in models of mechanochemical molecular machines—such as the Kawaguchi-Sasa-Sagawa (KSS) model for ATP synthase [134]—would be an interesting test-bed for both the applicability of the phenomenological expressions for excess power in such a model, and the relevance of the nonequilibrium contribution to the excess power under physiological conditions. Furthermore, the TSS excess work results obtained in Sec. 10.4.2 imply that the $\mathrm{F_o}$ component of ATP synthase may act as a fully autonomous Maxwell demon when controlling the dynamics of $\mathrm{F_1}$. It would be interesting to explore the implications of this with regard to thermodynamic performance metrics, such as efficiency [64], and whether or not other biological molecular machines admit similar thermodynamic interpretations.

Alternatively, instead of separating the entropy production of a mechanochemical system into contributions that are due to either the chemical or mechanical dynamics, a potentially more appealing decomposition of the total entropy production is to separate the so called visible (or coarse-grained) entropy production from the hidden entropy production [204].[1] This has been an active area of research in the stochastic thermodynamics community over the past several years [191, 204, 226, 227, 228, 229, 230, 231, 232], and addresses a number of practical issues with inferring entropy production in such systems, namely the differentiation between dynamics which are visible and dynamics which are not. However, the appearance of the generalized friction coefficient in quantifying the excess power between components of a strongly coupled system has promising implications for the use of similar phenomenological methods in approximating hidden entropy production.

[1] There are subtle, but important differences between splitting the entropy production into that which is due to each subsystem's dynamics, and that which is due to visible and hidden coordinates, even if one of the systems is visible and the other is hidden. Ultimately, this is because the transition rates of the visible system depend—through local detailed balance—on the state of the hidden system.

11.2 Final remark

Broadly, this thesis has presented a series of results which aim to further our understanding of the microscopic physics of strongly fluctuating systems. Ultimately, our aims have been to better elucidate and quantify the functional limitations and fundamental design principles of biological molecular machines. While any comprehensive understanding of such limitations is still in its infancy, we hope that the methods and techniques developed throughout this thesis are of use to further developments in this area. A distinguishing feature of our approach is the use of phenomenological coefficients obtained by linear-response theory or small-perturbation expansions—that can be obtained by measuring equilibrium fluctuations—to quantify nonequilibrium dissipation. In particular, our incorporation of the generalized friction coefficient into the hidden excess power framework serves to show the versatility of such approaches, applying broadly throughout the thesis, from experiments in Part I to the internal energy flows in Part III.

When one views the generalized friction tensor as a mathematical object giving geometric interpretations to the nonequilibrium physics of fluctuating systems, this continues a coherent arc in the study of thermodynamic systems that hearkens back to Gibbs. The "*general graphical method*" sought by Gibbs [119] to provide tractable interpretations of an otherwise abstract theory is continued in our work. Our ultimate goal is similar in nature to Gibbs': we hope that such simple theories describing the flows of energy, entropy, and information within and throughout nonequilibrium fluctuating systems will serve as a tool to build strong intuition, guiding the future investigation and understanding of molecular machines.

Bibliography

[1] L. Boltzmann, "The second law of thermodynamics," in *Theoretical physics and philosophical problems* (B. McGuinness, ed.), vol. 5, Springer, 1974.

[2] S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres a développer cette puissance*. Bachelier Libraire, 1824.

[3] D. V. Schroeder, *An Introduction to Thermal Physics*. Pearson, 1999.

[4] R. Clausius, "Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie," *Annalen der physik und chemie*, p. 23, 1865.

[5] R. Zwanzig, *Nonequilibrium Statistical Mechanics*. Oxford University Press, 2001.

[6] C. Jarzynski, "Nonequilibrium equality for free energy differences," *Phys. Rev. Lett.*, vol. 78, p. 2690, 1997.

[7] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," *Phys. Rev. E*, vol. 60, p. 2721, 1999.

[8] G. Gallavotti and E. G. D. Cohen, "Dynamical ensembles in nonequilibrium statistical mechanics," *Phys. Rev. Lett.*, vol. 74, p. 2694, 1995.

[9] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, "Probability of second law violations in shearing steady states," *Phys. Rev. Lett.*, vol. 71, p. 2401, 1993.

[10] D. J. Evans and D. J. Searles, "Equilibrium microstates which generate second law violating steady states," *Phys. Rev. E*, vol. 50, p. 1645, 1994.

[11] T. Hatano and S.-i. Sasa, "Steady-state thermodynamics of Langevin systems," *Phys. Rev. Lett.*, vol. 86, p. 3463, 2001.

[12] U. Seifert, "Entropy production along a stochastic trajectory and an integral fluctuation theorem," *Phys. Rev. Lett.*, vol. 95, p. 040602, 2005.

[13] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, "Efficiency at maximum power of low-dissipation Carnot engines," *Phys. Rev. Lett.*, vol. 105, p. 150603, 2010.

[14] J. M. Horowitz and H. Sandberg, "Second-law-like inequalities with information and their interpretations," *New J. Phys.*, vol. 16, p. 125007, 2014.

[15] T. Sagawa and M. Ueda, "Generalized Jarzynski equality under nonequilibrium feedback control," *Phys. Rev. Lett.*, vol. 104, p. 090602, 2010.

[16] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, "Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality," *Nat. Phys*, vol. 6, p. 988, 2010.

[17] K. Maruyama, F. Nori, and V. Vedral, "Colloquium: The physics of Maxwell's demon and information," *Rev. Mod. Phys.*, vol. 81, pp. 1–17, 2009.

[18] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. Res. Dev.*, vol. 5, p. 183, 1961.

[19] S. Ciliberto, "Experiments in stochastic thermodynamics: Short history and perspectives," *Phys. Rev. X*, vol. 7, p. 021051, 2017.

[20] J. R. Gomez-Solano, L. Bellon, A. Petrosyan, and S. Ciliberto, "Steady-state fluctuation relations for systems driven by an external random force," *EPL*, vol. 89, p. 60003, 2010.

[21] F. Douarche, S. Joubaud, N. B. Garnier, A. Petrosyan, and S. Ciliberto, "Work fluctuation theorems for harmonic oscillators," *Phys. Rev. Lett.*, vol. 97, p. 140603, 2006.

[22] N. Garnier and S. Ciliberto, "Nonequilibrium fluctuations in a resistor," *Phys. Rev. E*, vol. 71, p. 060101(R), 2004.

[23] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr., and C. Bustamante, "Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality," *Science*, vol. 296, pp. 1832–1835, 2002.

[24] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, "Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies," *Nature*, vol. 437, pp. 231–234, 2005.

[25] S. Tafoya, S. J. Large, S. Liu, C. Bustamante, and D. A. Sivak, "Using a system's equilibrium behaviour to reduce its energy dissipation in nonequilibrium processes," *Proc. Natl. Acad. Sci. USA*, vol. 116, pp. 5920–5924, 2019.

[26] E. Schrödinger, *What is Life?* Cambridge University Press, 1944.

[27] U. Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," *Rep. Prog. Phys.*, vol. 75, p. 126001, 2012.

[28] A. I. Brown and D. A. Sivak, "Theory of nonequilibrium free energy transduction by molecular machines," *Chem. Rev.*, vol. 120, p. 434, 2020.

[29] A. C. Barato and U. Seifert, "Thermodynamic uncertainty relation for biomolecular processes," *Phys. Rev. Lett.*, vol. 114, p. 158101, 2015.

[30] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, "Dissipation bounds all steady-state current fluctuations," *Phys. Rev. Lett.*, vol. 116, p. 120601, 2016.

[31] P. Pietzonka, F. Ritort, and U. Seifert, "Finite-time generalization of the thermodynamic uncertainty relation," *Phys. Rev. E*, vol. 96, p. 012101, 2017.

[32] T. Koyuk, U. Seifert, and P. Pietzonka, "A generalization of the thermodynamic uncertainty relation to periodically driven systems," *J. Phys. A: Math. Theor.*, vol. 52, p. 02LT02, 2019.

[33] J. M. Horowitz and T. R. Gingrich, "Proof of finite-time thermodynamic uncertainty relation for steady-state currents," *Phys. Rev. E*, vol. 96, p. 020103(R), 2017.

[34] J. Horowitz and T. R. Gingrich, "Thermodynamic uncertainty relations constrain non-equilibrium fluctutations," *Nat. Phys.*, vol. 16, pp. 15–20, 2020.

[35] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton*. Sinauer Associates, 2001.

[36] K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block, "Direct observation of kinesin stepping by optical trapping interferometry," *Nature*, vol. 365, pp. 721–727, 1993.

[37] K. Svoboda and S. M. Block, "Force and velocity measured for single kinesin molecules," *Cell*, vol. 77, pp. 773–784, 1994.

[38] H. Kojima, E. Muto, H. Higuchi, and T. Yanagida, "Mechanics of single kinesin molecules measured by optical trapping nanometry," *Biophys. J.*, vol. 73, pp. 2012–2022, 1997.

[39] A. J. Hunt, F. Gittes, and J. Howard, "The force exerted by a single kinesin molecule against a viscous load," *Biophys. J.*, vol. 67, pp. 766–781, 1994.

[40] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, "Nonequilibrium energetics of a single F_1-ATPase molecule," *Phys. Rev. Lett.*, vol. 104, p. 198103, 2010.

[41] M. Yoshida, E. Muneyuki, and T. Hisabori, "ATP synthase – a marvellous rotary engine of the cell," *Nat. Rev. Mol. Cell Biol.*, vol. 2, p. 669, 2001.

[42] H. Ueno, T. Suzuki, K. Kinosita Jr., and M. Yoshida, "ATP-driven stepwise rotation of F_oF_1-ATP synthase," *Proc. Natl. Acad. Sci. USA*, vol. 105, pp. 1333–1338, 2005.

[43] J. Chen, A. Coakley, M. O'Connor, A. Petrov, S. E. O'Leary, J. F. Atkins, and J. D. Puglisi, "Coupling of mRNA structure rearrangement to ribosome movement during bypassing of noncoding regions," *Cell*, vol. 163, pp. 1267–1280, 2015.

[44] H. Peng, X.-F. Li, H. Zhang, and C. Le, "A microRNA-initiated DNAzyme motor operating in living cells," *Nat. Comm.*, vol. 8, p. 14378, 2017.

[45] R. Phillips, J. Kondev, J. Theriot, and H. Garcia, *Physical Biology of the Cell*. Garland Science, 2nd ed., 2012.

[46] S. A. Endow, F. J. Kull, and H. Liu, "Kinesins at a glance," *J. Cell. Sci.*, vol. 123, p. 3420, 2010.

[47] S. T. Brady, R. J. Lasek, and S. T. Brady, "Fast axonal transport in extruded axoplasm from squid giant axon," *Science*, vol. 218, no. 4577, pp. 1129–1131, 1982.

[48] R. D. Allen, J. Metuzals, I. Tasaki, S. T. Brady, and S. P. Gilbert, "Fast axonal transport in squid giant axon," *Science*, vol. 218, no. 4577, pp. 1127–1129, 1982.

[49] R. D. Vale, T. S. Reese, and M. P. Sheetz, "Identification of a novel force-generating protein, kinesin, involved in mircotubule-based motility," *Cell*, vol. 42, no. 1, pp. 39–50, 1985.

[50] J. T. Yang, R. A. Laymon, and L. S. B. Goldstein, "A three-domain structure of kinesin heavy chain revealed by DNA sequence and microtubule binding analyses," *Cell*, vol. 56, pp. 879–889, 1989.

[51] G. Karp, J. Iwasa, and W. Marshall, *Karp's Cell and Molecular Biology: Concepts and Experiments*. Wiley, 8th ed., 2016.

[52] E. P. Sablin, R. B. Case, S. C. Dai, C. L. Hart, A. Ruby, R. D. Vale, and R. J. Fletterick, "Direction determination in the minus-end-directed kinesin motor ncd," *Nature*, vol. 395, pp. 813–816, 1998.

[53] R. A. Cross, "Molecular motors: The natural economy of kinesin," *Curr. Biol.*, vol. 7, pp. R631–R633, 1997.

[54] D. Cai, K. J. Verhey, and E. Meyhöfer, "Tracking single kinesin molecules in the cytoplasm of mammalian cells," *Biopys. J.*, vol. 92, pp. 4137–4144, 2007.

[55] L. S. B. Goldstein, "Kinesin molecular motors: Transport pathways, receptors, and human disease," *Proc. Natl. Acad. Sci. USA*, vol. 98, no. 13, pp. 6999–7003, 2001.

[56] W. Guo, K. S. Dittlau, and L. V. D. Bosch, "Axonal transport defects and neurodegeneration: Molecular mechanisms and therapeutic implications," *Semin. Cell Dev. Biol.*, vol. 99, pp. 133–150, 2020.

[57] M. Yoshida, E. Muneyuki, and T. Hisabori, "ATP synthase – a marvelous rotary engine of the cell," *Nat. Rev. Cell. Mol. Bio.*, vol. 2, p. 669, 2001.

[58] P. D. Boyer, "The ATP synthase – a splendid molecular machine," *Annu. Rev. Biochem.*, vol. 66, pp. 717–49, 1997.

[59] M. Lynch and G. K. Marinov, "Membranes, energetics, and evolution across the prokaryote-eukaryote divide," *eLife*, vol. 6, p. e20437, 2017.

[60] S. Pamarthy, A. Kulshrestha, G. K. Katara, and K. D. Beaman, "The curious case of vacuolar ATPase: regulation of signaling pathways," *Mol. Cancer*, vol. 17, p. 41, 2018.

[61] M. Forgac, "Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology," *Nat. Rev. Mol. Cell Bio.*, vol. 8, pp. 917–929, 2007.

[62] R. L. Cross and V. Müller, "The evolution of A- F- and V-type ATP synthases and ATPases: reversals in function and changes in the H^+/ATP coupling ratio," *FEBS Lett.*, vol. 576, pp. 1–4, 2004.

[63] T. P. Silverstein, "An exploration of how the thermodynamic efficiency of bioenergetic membrane systems arises with *c*-subunit stoichiometry of F_1F_oATP synthases," *J. Bioenerg. Biomembr.*, vol. 46, p. 229, 2014.

[64] E. Lathouwers, J. N. E. Lucero, and D. A. Sivak, "Nonequilibrium energy transduction in stochastic strongly coupled rotary motors," *J. Phys. Chem. Lett.*, vol. 11, p. 5273, 2020.

[65] J. N. E. Lucero, A. Mehdizadeh, and D. A. Sivak, "Optimal control of rotary motors," *Phys. Rev. E*, vol. 99, p. 012119, 2019.

[66] A. K. S. Kasper and D. A. Sivak, "Modeling work-speed-accuracy trade-offs in a stochastic rotary machine," *Phys. Rev. E*, vol. 101, p. 032110, 2020.

[67] A. B. Kolomeisky and M. E. Fisher, "Molecular motors: a theorist's perspective," *Annu. Rev. Phys. Chem.*, vol. 58, p. 675, 2007.

[68] M. Esposito, "Stochastic thermodynamics under coarse graining," *Phys. Rev. E*, vol. 85, p. 041125, 2012.

[69] R. Brown, "A brief account of microscopical observations made on the particles contained in the pollen of plants," *Phil. Mag.*, vol. 4, p. 121, 1828.

[70] A. Einstein, "Über die von der molekular-knetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen," *Ann. Phys.*, vol. 17, p. 549, 1905.

[71] M. von Smoluchowski, "Zur kinetischen theorie der brownsche bewegung," *Ann. Phys.*, vol. 21, p. 756, 1906.

[72] C. Gardiner, *Stochastic Methods, A Handbook for the Natural and Social Sciences*. Springer, 4th ed., 2009.

[73] K. Du, J. A. Liddle, and A. J. Berglund, "Three-dimensional real-time tracking of nanoparticles at an oil–water interface," *Langmuir*, vol. 28, pp. 9181–9188, 2012.

[74] B. Radoev, M. Nedjalkov, and V. Djakovich, "Brownian motion at liquid-gas interfaces. 1. Diffusion coefficients of macroparticles at pure interfaces," *Langmuir*, vol. 8, pp. 2962–2965, 1992.

[75] J. Toro-Mendoza, G. Rodriguez-Lopez, and O. Paredes-Altuve, "Brownian diffusion of a particle at an air/liquid interface: the elastic (not viscous) response of the surface," *Phys. Chem. Chem. Phys.*, vol. 19, p. 9092, 2017.

[76] S. J. Large, R. Chetrite, and D. A. Sivak, "Stochastic control in microscopic nonequilibrium systems," *EPL*, vol. 124, p. 20001, 2018.

[77] S. J. Large and D. A. Sivak, "Optimal discrete control: minimizing dissipation in discretely driven nonequilibrium systems," *J. Stat. Mech.*, p. 083212, 2019.

[78] S. J. Large, J. Ehrich, and D. A. Sivak, "Free energy transduction within autonomous machines." 2020, arxiv:2008.03402.

[79] S. J. Large and D. A. Sivak, "Hidden energy flows in strongly coupled nonequilibrium systems," *EPL*, 2020, in press, arxiv:2009.07973.

[80] A. Kolmogorov, *Foundations of the Theory of Probability*. Chelsea publishing company, 1950.

[81] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems*. Cambridge University Press, 2010.

[82] D. W. Stroock, *Probability Theory: an Analytic View*. Cambridge University Press, 2010.

[83] L. A. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*. Springer, 2013.

[84] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*. Elsevier, 3rd ed., 2007.

[85] I. N. Sneddon, *Fourier Transforms*. Dover Publications, 2010.

[86] H. Touchette, "The large deviation approach to statistical mechanics," *Phys. Rep.*, vol. 478, pp. 1–69, 2009.

[87] D. Williams, *Probability with Martingales*. Cambridge Mathematical Textbooks, Cambridge University Press, 1991.

[88] R. Rao and M. Esposito, "Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics," *Phys. Rev. X*, vol. 6, p. 041064, 2016.

[89] H. Risken, *The Fokker-Planck equation: Methods of solution and applications*. Springer-Verlag, 2nd ed., 2011.

[90] L. D. Landau and E. M. Lifshitz, *Fluid Dynamics*. Pergamon Press, 2nd ed., 1987.

[91] E. M. Purcell, "Life at low Reynolds number," *Am. J. Phys.*, vol. 45, pp. 3–11, 1977.

[92] K. Sekimoto, "Langevin equations and thermodynamics," *Prog. Theor. Phys. Supp.*, vol. 130, p. 17, 1998.

[93] C. Jarzynski, "Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale," *Annu. Rev. Condens. Matter. Phys.*, vol. 2, pp. 329–351, 2011.

[94] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2nd ed., 2006.

[95] G. N. Bochkov and Y. E. Kuzovlev, "General theory of thermal fluctuations in nonlinear systems," *Sov. Phys. JETP*, vol. 45, p. 125, 1977.

[96] G. N. Bochkov and Y. E. Kuzovlev, "Fluctuation-dissipation relations for nonequilibrium processes in open systems," *Sov. Phys. JETP*, vol. 49, p. 543, 1978.

[97] M. Esposito and C. Van den Broeck, "Three detailed fluctuation theorems," *Phys. Rev. Lett.*, vol. 104, p. 090601, 2010.

[98] M. Esposito and C. Van den Broeck, "Three faces of the second law. I. Master equation formulation," *Phys. Rev. E*, vol. 82, p. 011143, 2010.

[99] C. Van den Broeck and M. Esposito, "Three faces of the second law. II. Fokker-Planck formulation," *Phys. Rev. E*, vol. 82, p. 011144, 2010.

[100] G. E. Crooks and D. A. Sivak, "Measures of trajectory ensemble disparity in nonequilibrium statistical dynamics," *J. Stat. Mech.*, p. P06003, 2011.

[101] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamical Systems*. Prentice-Hall, 4th ed., 2002.

[102] J. C. Maxwell, "On governors," *Proc. R. Soc. London*, vol. 16, pp. 270–283, 1868.

[103] J. Bechhoefer, "Feedback for physicists: A tutorial essay on control," *Rev. Mod. Phys.*, vol. 77, p. 783, 2005.

[104] K. Proesmans, J. Ehrich, and J. Bechhoefer, "Finite-time Landauer principle," *Phys. Rev. Lett.*, vol. 125, p. 100602, 2020.

[105] K. Proesmans, J. Ehrich, and J. Bechhoefer, "Optimal finite-time bit erasure under full control," *Phys. Rev. E*, vol. 102, p. 032105, 2020.

[106] T. Schmiedl and U. Seifert, "Optimal finite-time processes in stochastic thermodynamics," *Phys. Rev. Lett.*, vol. 98, p. 108301, 2007.

[107] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, "Optimal protocols and optimal transport in stochastic thermodynamics," *Phys. Rev. Lett.*, vol. 106, p. 250601, 2011.

[108] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, "Boundary layers in stochastic thermodynamics," *Phys. Rev. E*, vol. 85, p. 020103(R), 2012.

[109] D. A. Sivak and G. E. Crooks, "Thermodynamic metrics and optimal paths," *Phys. Rev. Lett.*, vol. 108, p. 190602, 2012.

[110] D. Chandler, *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.

[111] N. G. van Kampen, "The case against linear response theory," *Phys. Norv.*, vol. 5, p. 279, 1971.

[112] A. Sarracino and A. Vulpiani, "On the fluctuation-dissipation relation in non-equilibrium and non-Hamiltonian systems," *Chaos*, vol. 29, p. 083132, 2019.

[113] W. J. Rugh, *Nonlinear System Theory: The Volterra-Wiener Approach*. Johns Hopkins University Press, 1981.

[114] J. G. Kirkwood, "The statistical mechanical theory of transport processes I. General theory," *J. Chem. Phys.*, vol. 14, p. 180, 1964.

[115] R. Zwanzig, "Elementary derivation of time-correlation formulas for transport coefficients," *J. Chem. Phys.*, vol. 40, p. 2527, 1964.

[116] D. A. Garanin, "Integral relaxation time of single-domain ferromagnetic particles," *Phys. Rev. E*, vol. 54, p. 3250, 1996.

[117] P. R. Zulkowski and M. R. DeWeese, "Optimal finite-time erasure of a classical bit," *Phys. Rev. E*, vol. 89, p. 052140, 2014.

[118] A. Berezhkovskii and A. Szabo, "Time scale separation leads to position-dependent diffusion along a slow coordinate," *J. Chem. Phys.*, vol. 135, p. 174108, 2011.

[119] M. J. Klein, "The physics of J. Willard Gibbs in his time," *Phys. Today*, vol. 43, no. 3, p. 40, 1990.

[120] F. Weinhold, "Metric geometry of equilibrium thermodynamics," *J. Chem. Phys.*, vol. 63, no. 6, pp. 2479–2483, 1975.

[121] G. Ruppeiner, "Thermodynamics: A Riemannian geometric model," *Phys. Rev. A*, vol. 20, p. 1608, 1979.

[122] F. Schlögl, "A connection between correlations and the order of bit-number cumulants," *Z. Phys. B*, vol. 59, p. 449, 1985.

[123] J. Nulton, P. Salamon, B. Andresen, and Q. Amin, "Quasistatic processes as step equilibrations," *J. Chem. Phys.*, vol. 83, p. 334, 1985.

[124] G. E. Crooks, "Measuring thermodynamic length," *Phys. Rev. Lett.*, vol. 99, p. 100602, 2007.

[125] G. M. Rotskoff and G. E. Crooks, "Optimal control in nonequilibrium systems: Dynamic Riemannian geometry of the Ising model," *Phys. Rev. E*, vol. 92, p. 060102(R), 2015.

[126] P. R. Zulkowski and M. R. DeWeese, "Optimal control of overdamped systems," *Phys. Rev. E*, vol. 92, p. 032117, 2015.

[127] D. A. Sivak and G. E. Crooks, "Thermodynamic geometry of minimum-dissipation driven barrier crossing," *Phys. Rev. E*, vol. 94, p. 052106, 2016.

[128] P. R. Zulkowski, D. A. Sivak, G. E. Crooks, and M. R. DeWeese, "Geometry of thermodynamic control," *Phys. Rev. E*, vol. 86, p. 041148, 2012.

[129] J. E. Niven and S. B. Laughlin, "Energy limitation as a selective pressure on the evolution of sensory systems," *J. Exp. Biol.*, vol. 211, p. 1792, 2008.

[130] A. P. Solon and J. M. Horowitz, "Phase transitions in protocols minimizing work fluctuations," *Phys. Rev. Lett.*, vol. 120, p. 180605, 2018.

[131] O. Mazonka and C. Jarzynski, "Exactly solvable model illustrating far-from-equilibrium predictions." 1999, arxiv:9912121.

[132] P. Reimann, "Brownian motors: noisy transport far from equilibrium," *Phys. Rep.*, vol. 361, p. 57, 2002.

[133] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2nd ed., 2002.

[134] K. Kawguchi, S.-i. Sasa, and T. Sagawa, "Nonequilibrium dissipation-free transport in F_1-ATPase and the thermodynamic role of asymmetric allosterism," *Biophys. J.*, vol. 106, p. 2450, 2014.

[135] S. Toyabe, H. Ueno, and E. Muneyuki, "Recovery of state-specific potential of molecular motor from single-molecule trajectory," *EPL*, vol. 97, p. 40004, 2012.

[136] H. B. Callen, *Thermodynamics and an Introduction to Thermostatics*. Wiley, 1985.

[137] Z. T. Berndsen, N. Keller, S. Grimes, P. J. Jardine, and D. E. Smith, "Nonequilibrium dynamics and ultraslow relaxation of confined DNA during viral packaging," *Proc. Natl. Acad. Sci. USA*, vol. 111, pp. 8345–8350, 2014.

[138] M. Besprozvannaya, V. L. Pivorunas, Z. Feldman, and B. M. Burton, "SpoIIIE protein achieves directional DNA translocation through allosteric regulation of ATPase activity by an accessory domain," *J. Biol. Chem.*, vol. 288, pp. 28962–28974, 2013.

[139] K. K. Jr., R. Yasuda, H. Noji, and K. Adachi, "A rotary molecular motor that can work at near 100% efficiency," *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, vol. 355, pp. 473–489, 2000.

[140] D. Bikard, C. Loot, Z. Baharoglu, and D. Mazel, "Folded DNA in action: Hairpin formation and biological functions in prokaryotes," *Mirobiol. Mol. Biol. R.*, vol. 74, pp. 570–588, 2010.

[141] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, "Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins," *Proc. Natl. Acad. Sci. USA*, vol. 103, pp. 6190–6195, 2006.

[142] M. A. Glucksmann-Kuis, X. Dai, P. Markiewicz, and L. B. Rothman-Denes, "E. coli SSB activates N4 virion RNA polymerase promoters by stabilizing a DNA hairpin required for promoter recognition," *Cell*, vol. 84, pp. 147–154, 1996.

[143] D. Wang, T. I. Meier, C. L. Chan, G. Feng, D. N. Lee, and R. Landick, "Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex," *Cell*, vol. 81, pp. 341–350, 1995.

[144] T. Liu, A. Kaplan, L. Alexander, S. Yan, J.-D. Wen, L. L. Jr, C. E. Wickersham, K. Fredrick, H. Noller, I. T. Jr, and C. J. Bustamante, "Direct measurement of the mechanical work during translocation by the ribosome," *eLife*, vol. 3, p. e03406, 2014.

[145] L. J. Simpson, E. Tzima, and J. S. Reader, "Mechanical force and their effect on the ribosome and protein translation machinery," *Cells*, vol. 9, p. 650, 2020.

[146] G. Varani, "Exceptionally stable nucleic acid hairpins," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 24, pp. 379–404, 1995.

[147] S. J. Green, D. Lubrich, and A. J. Turberfield, "DNA hairpins: Fuel for autonomous DNA devices," *Biophys. J.*, vol. 91, pp. 2966–2975, 2006.

[148] A. Eshra, S. Shah, T. Song, and J. Reif, "Renewable DNA hairpin-based logic circuits," *IEEE T. Nanotechnol.*, vol. 18, p. 252, 2019.

[149] J. R. Moffitt, Y. R. Chemla, S. B. Smith, and C. Bustamante, "Recent advances in optical tweezers," *Annu. Rev. Biochem.*, vol. 77, pp. 205–228, 2008.

[150] F. Ritort, "Single-molecule experiments in biological physics:methods and applications," *J. Phys.: Condens. Matter*, vol. 18, pp. R531–R538, 2006.

[151] M. Bercy and U. Bockelmann, "Hairpins under tension: RNA versus DNA," *Nucleic Acids Res.*, vol. 43, pp. 9928–9936, 2015.

[152] C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith, "Single-molecular studies of DNA mechanics," *Curr. Opin. Struc. Biol.*, vol. 10, pp. 279–285, 2000.

[153] K. Neupane, F. Wang, and M. T. Woodside, "Direct measurement of sequence-dependent transition path times and conformational diffusion in DNA duplex formation," *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 6, p. 1329, 2017.

[154] D. A. N. Foster, R. Petrosyan, A. G. T. Pyo, A. Hoffman, F. Wang, and M. T. Woodside, "Probing position-dependent diffusion in folding reactions using single-molecule force spectroscopy," *Biophys. J.*, vol. 114, p. 1657, 2018.

[155] C. Bustamante, Y. R. Chemla, and J. R. Moffitt, "High-resolution dual-trap optical tweezers with differential detection: Instrument design," *Cold Spring Harb. Protoc.*, 2009.

[156] C. Bustamante, Y. R. Chemla, and J. R. Moffitt, "High-resolution dual-trap optical tweezers with differential detection: Data collection and instrument calibration," *Cold Spring Harb. Protoc.*, 2009.

[157] P. T. X. Li, C. Bustamante, and I. Tinoco Jr., "Real-time control of the energy landscape by force directs the folding of RNA molecules," *Proc. Natl. Acad. Sci. USA*, vol. 104, pp. 7039–7044, 2007.

[158] C. N. H. Candia, S. T. Martinez, and B. Guiérrez-Medina, "A minimal optical trapping and imaging microscopy system," *PLoS One*, vol. 8, p. e57383, 2013.

[159] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and their Application*. Cambridge University Press, 1997.

[160] T. Hastie, R. Tibshirani, and J. Friedman, *Elements of Statistical Learning*. Springer-Verlag, 2016.

[161] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. 19, pp. 716–723, 1974.

[162] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco Jr., and C. Bustamante, "Reversible unfolding of single RNA molecules by mechanical force," *Science*, vol. 292, pp. 733–737, 2001.

[163] A. Alemany and F. Ritort, "Force-dependent folding and unfolding kinetics in DNA hairpins reveals transition-state displacements along a single pathway," *J. Phys. Chem. Lett.*, vol. 8, pp. 895–900, 2017.

[164] O. K. Dudko, G. Hummer, and A. Szabo, "Intrinsic rates and activation free energies from single-molecule pulling experiments," *Phys. Rev. Lett.*, vol. 96, p. 108101, 2006.

[165] L. Seoane and R. Solé, "Phase transitions in Pareto optimal complex networks," *Phys. Rev. E*, vol. 92, p. 032807, 2015.

[166] Z.-J. Tan and S.-J. Chen, "Nucleic acid helix stability: Effects of salt concentration, cation valence and size, and chain length," *Biophys. J.*, vol. 90, p. 1175, 2006.

[167] Z.-J. Tan and S.-J. Chen, "Salt dependence of nucleic acid hairpin stability," *Biophys. J.*, vol. 95, p. 738, 2008.

[168] M.-N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette, "Stretching single stranded DNA, a model polyelectrolyte," *Phys. Rev. Lett.*, vol. 89, p. 248102, 2002.

[169] A. Bosco, J. Camunas-Soler, and F. Ritort, "Elastic properties and secondary structure formation of single-stranded DNA at monovalent an divalent salt concentrations," *Nucleic Acids Res.*, vol. 42, p. 2064, 2013.

[170] C. M. H. Ferreira, I. S. S. Pinto, E. V. Soares, and H. M. V. M. Soares, "(un)suitability of the use of pH buffers in biological biochemical and environmental studies and their interaction with metal ions – a review," *RSC Adv.*, vol. 5, p. 30989, 2015.

[171] G. M. Rotskoff and P. L. Geissler, "Robust pathways to multicompartment assembly," *Proc. Natl. Acad. Sci. USA*, vol. 115, no. 25, pp. 6341–6346, 2018.

[172] B. Israelewitz, M. Gao, and K. Schulten, "Steered molecular dynamics and mechanical function of proteins," *Curr. Opin. Struct. Biol.*, vol. 11, pp. 224–230, 2001.

[173] J. Gore, F. Ritort, and C. Bustamante, "Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements," *Proc. Natl. Acad. Sci. USA*, vol. 100, pp. 12564–12569, 2003.

[174] S. Liu, G. Chistol, C. L. Hetherington, S. Tafoya, K. Aathavan, J. Schnitzbauer, S. Grimes, P. J. Jardine, and C. Bustamante, "A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills," *Cell*, vol. 157, pp. 703–713, 2014.

[175] V. P. Desai, F. Frank, A. Lee, M. Righini, L. Lancaster, H. F. Noller, I. Tinoco Jr., and C. Bustamante, "Co-temporal force and fluorescence measurements reveal a ribosomal gear shift mechanism of translation regulation by structured mRNAs," *Mol. Cell*, vol. 75, pp. 1007–1019, 2019.

[176] R. Yasuda, H. Noji, Kazuhiko Kinosita Jr., and M. Yoshida, "F_1-ATPase is a highly efficient molecular motor that rotates with discrete 120^o steps," *Cell*, vol. 93, pp. 1117–1124, 1998.

[177] H. Itoh, A. Takahashi, K. Adachi, H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita Jr., "Mechanically driven ATP synthesis by F_1-ATPase," *Nature*, vol. 427, pp. 465–468, 2004.

[178] Y. Rondelez, G. Tresset, T. Nakashima, Y. Kato-Yamada, H. Fujita, S. Takeuchi, and H. Noji, "Highly coupled ATP synthesis by F_1-ATPase single molecules," *Nature*, vol. 433, pp. 773–777, 2005.

[179] G. Casella and R. R. Berger, *Statistical Inference*. Thompson Learning, 2002.

[180] B. B. Machta, "Dissipation bound for thermodynamic control," *Phys. Rev. Lett.*, vol. 115, p. 260603, 2015.

[181] J. M. Horowitz and M. Esposito, "Work producing reservoirs: Stochastic thermodynamics with generalized Gibbs ensembles," *Phys. Rev. E*, vol. 94, p. 020102(R), 2016.

[182] A. C. Barato and U. Seifert, "Thermodynamic cost of external control," *New J. Phys.*, vol. 19, p. 073021, 2017.

[183] G. Verley, C. Van den Broeck, and M. Esposito, "Work statistics of stochastically driven systems," *New J. Phys.*, vol. 16, p. 095001, 2014.

[184] A. Gomez-Marin, T. Schmiedl, and U. Seifert, "Optimal protocols for minimal work processes in underdamped stochastic thermodynamics," *J. Chem. Phys.*, vol. 129, p. 024114, 2008.

[185] J. Burbea and C. R. Rao, "Entropy differential metric, distance and divergence measures in probability spaces: A unified approach," *J. Multivariate Anal.*, vol. 12, pp. 575–596, 1982.

[186] B. Andreson, P. Salamon, and R. S. Berry, "Thermodynamics in finite time," *Phys. Today*, vol. 37, p. 62, 1984.

[187] P. Salamon and R. S. Berry, "Thermodynamic length and dissipated availability," *Phys. Rev. Lett.*, vol. 51, p. 1127, 1983.

[188] P. Salamon, J. Nulton, and E. Ihrig, "On the relation between entropy and energy versions of thermodynamic length," *J. Chem. Phys.*, vol. 80, p. 436, 1984.

[189] W. L. Winston, *Operations Research: Applications and Algorithms*. Duxbury Press, 1978.

[190] S. J. Bryant and B. B. Machta, "Energy dissipation bounds for autonomous thermodynamic cycles," *Proc. Natl. Acad. Sci. USA*, vol. 117, no. 7, pp. 3478–3483, 2020.

[191] I. A. Martínez, G. Basker, J. M. Horowitz, and J. M. R. Parrondo, "Inferring broken detailed balance in the absence of observable currents," *Nat. Comm.*, vol. 10, p. 3542, 2019.

[192] D. L. Floyd, S. C. Harrison, and A. M. van Oijen, "Analysis of kinetic intermediates in single-particle dwell-time distributions," *Biophys. J.*, vol. 99, pp. 360–366, 2010.

[193] G. E. Crooks, *A Field Guide to Continuous Probability Distributions*. Berkeley Institute for Theoretical Science, 2019.

[194] D. V. Schroeder, *An Introduction to Thermal Physics*. Pearson, 2000.

[195] J. M. Horowitz and M. Esposito, "Thermodynamics with continuous information flow," *Phys. Rev. X*, vol. 4, p. 031015, 2014.

[196] J. Wagoner and K. Dill, "Mechanisms for achieving high speed and efficiency in biomolecular machines," *Proc. Natl. Acad. Sci. USA*, vol. 116, p. 5902, 2019.

[197] S. Toyabe, T. Watanabe-Nakayama, T. Okamoto, S. Kudo, and E. Muneyuki, "Thermodynamic efficiency and mechanochemical coupling of F_1-ATPase," *Proc. Natl. Acad. Sci. USA*, vol. 108, p. 17951, 2011.

[198] W. Li and A. Ma, "Reaction mechanism and reaction coordinates from the viewpoint of energy flow," *J. Chem. Phys.*, vol. 114, p. 114103, 2016.

[199] D. Hartich, A. C. Barato, and U. Seifert, "Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell's demon interpretation," *J. Stat. Mech.*, p. P02016, 2014.

[200] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, "Thermodynamics of prediction," *Phys. Rev. Lett.*, vol. 109, p. 120604, 2012.

[201] M. E. Quenneville and D. A. Sivak, "Energy dissipation and information flow in coupled Markovian systems," *entropy*, vol. 20, p. 707, 2018.

[202] A. Barato, D. Hartich, and U. Seifert, "Efficiency of cellular information processing," *New J. Phys.*, vol. 16, p. 103024, 2014.

[203] R. A. Brittain, N. S. Jones, and T. E. Ouldridge, "What we learn from the learning rate," *J. Stat. Mech.*, vol. 2017, p. 063502, 2017.

[204] U. Seifert, "From stochastic thermodynamics to thermodynamic inference," *Annu. Rev. Condens. Matter Phys.*, vol. 10, pp. 171–92, 2019.

[205] C. Van den Broeck and M. Esposito, "Ensemble and trajectory thermodynamics: A brief introduction," *Physica A*, vol. 418, pp. 6–16, 2015.

[206] P. G. Bergmann and J. L. Lebowitz, "New approach to nonequilibrium processes," *Phys. Rev.*, vol. 99, p. 578, 1955.

[207] M. Esposito and C. Van den Broeck, "Second law and Landauer principle far from equilibrium," *EPL*, vol. 95, p. 40004, 2011.

[208] A. I. Brown and D. A. Sivak, "Allocating dissipation across a molecular machine cycle to maximize flux," *Proc. Natl. Acad. Sci. USA*, vol. 114, p. 11057, 2017.

[209] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics*. Springer, 2nd ed., 1998.

[210] N. Soga, K. Kimura, K. Kinosita Jr., M. Yoshida, and T. Suzuki, "Perfect chemo-mechanical coupling of F_oF_1-ATP synthase," *Proc. Natl. Acad. Sci. USA*, vol. 19, pp. 4960–4965, 2017.

[211] J. Xing, J.-C. Liao, and G. Oster, "Making ATP," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 46, pp. 16539–16546, 2005.

[212] H. Wang and G. Oster, "Ratchets, power strokes, and molecular motors," *Appl. Phys. A*, vol. 75, pp. 315–323, 2002.

[213] B. Altaner, A. Wachtel, and J. Vollmer, "Fluctuating currents in stochastic thermodynamics. ii. energy conversion and nonequilibrium response in kinesin models," *Phys. Rev. E*, vol. 92, p. 042133, 2015.

[214] M. T. Valentine, P. M. Fordyce, T. C. Krzysiak, S. P. Gilbert, and S. M. Block, "Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads *in vitro*," *Nat. Cell Biol.*, vol. 8, pp. 470–476, 2006.

[215] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, "Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales," *Phys. Rev. Lett.*, vol. 89, p. 050601, 2002.

[216] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.

[217] D. A. Sivak and G. E. Crooks, "Near-equilibrium measurements of nonequilibrium free energy," *Phys. Rev. Lett.*, vol. 108, p. 150601, 2012.

[218] T. Ariga, M. Tomishige, and D. Mizuno, "Nonequilibrium energetics of molecular motor Kinesin," *Phys. Rev. Lett.*, vol. 121, p. 218101, 2018.

[219] M. Ribezzi-Crivellari and F. Ritort, "Large work extraction and the Landauer limit in a continuous Maxwell demon," *Nat. Phys.*, vol. 15, pp. 660–664, 2019.

[220] M. Ribezzi-Crivellari and F. Ritort, "Work extraction, information-content and the Landauer bound in the continuous Maxwell Demon," *J. Stat. Mech.*, p. 084013, 2019.

[221] P. A. Camati, J. P. S. Peterson, T. B. B. ao, K. Micadei, and A. M. Souza, "Experimental rectification of entropy production my Maxwell's demon in a quantum system," *Phys. Rev. Lett.*, vol. 117, p. 240502, 2016.

[222] J. Koski, A. Kutvonen, I. M. Khaymovich, T. Ala-Nissila, and J. P. Pekola, "On-chop Maxwell's demon as an information-powered refridgerator," *Phys. Rev. Lett.*, vol. 115, p. 260602, 2015.

[223] T. K. Saha, J. N. E. Lucero, J. Ehrich, D. A. Sivak, and J. Bechhoefer, "Maximizing power and velocity of an information engine." 2020, arxiv:2011.05478.

[224] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, "Thermodynamics of information," *Nat. Phys.*, vol. 11, pp. 131–139, 2015.

[225] M. E. Cates, "Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics?," *Rep. Prog. Phys.*, vol. 75, p. 042601, 2012.

[226] J. Li, J. M. Horowitz, T. R. Gingrich, and N. Fakhri, "Quantifying dissipation using fluctuating currents," *Nat. Comm.*, vol. 10, p. 1666, 2019.

[227] Édgar Roldán and J. M. R. Parrondo, "Estimating dissipation from single stationary trajectories," *Phys. Rev. Lett.*, vol. 105, p. 150607, 2010.

[228] A. Alemany, M. Ribezzi-Crivellari, and F. Ritort, "From free energy measurements to thermodynamic inference in nonequilibrium small systems," *New J. Phys.*, vol. 17, p. 075009, 2015.

[229] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, "Role of hidden slow degrees of freedom in the fluctuation theorem," *Phys. Rev. Lett.*, vol. 108, p. 220601, 2012.

[230] M. Uhl, P. Pietzonka, and U. Seifert, "Fluctuations of apparent entropy production in networks with hidden slow degrees of freedom," *J. Stat. Mech.*, p. 023203, 2018.

[231] M. Kahlen and J. Ehrich, "Hidden slow degrees of freedom and fluctuation theorems: an analytically solvable model," *J. Stat. Mech.*, p. 063204, 2018.

[232] D. Gupta and S. Sabhapandit, "Entropy production for partially observed harmonic systems," *J. Stat. Mech.*, p. 013204, 2020.

[233] R. W. Pastor, B. R. Brooks, and A. Szabo, "An analysis of the accuracy of Langevin and molecular dynamics algorithms," *Mol. Phys.*, vol. 65, pp. 1409–1419, 1988.

[234] D. A. Sivak, J. D. Chodera, and G. E. Crooks, "Using nonequilibrium fluctuations theorems to understand and correct errors in equilibrium and nonequilibrium simulations of discrete Langevin dynamics," *Phys. Rev. X*, vol. 3, p. 011007, 2013.

[235] D. A. Sivak, J. D. Chodera, and G. E. Crooks, "Thime step rescaling recovers continuous-time dynamical properties for discrete-time Langevin integration of nonequilibrium systems," *J. Phys. Chem. B*, vol. 118, pp. 6466–6474, 2014.

[236] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rend. Circ. Mat. Palermo*, vol. 4, pp. 48–90, 1955.

[237] A. Borodin and P. Salminen, *Handbook of Brownian Motion - Facts And Formula.* Springer, 2000.

[238] E. Selletin, M. Quartin, and L. Amendola, "Breaking the spell of Gaussianity: forecasting with higher-order Fisher matrices," *Mon. Not. R. Astron. Soc.*, vol. 441, pp. 1831–1840, 2014.

[239] D. J. Bacon, D. M. Goldberg, B. T. P. Rowe, and A. N. Taylor, "Weak gravitational flexion," *Mon. Not. R. Astron. Soc.*, vol. 365, pp. 414–428, 2006.

[240] D. M. Goldberg and D. J. Bacon, "Galaxy-galaxy flexion: weak lensing to second order," *Astrophys. J.*, vol. 619, pp. 741–748, 2005.

Appendix A

Numerical details

This appendix presents details of the numerical calculations and simulations presented throughout this thesis. We have made use of both Langevin and Master equation simulations, and in the following two sections we will outline our approach to both systems.[1]

A.1 Master equation

The master equation 2.2.1, as used in this thesis, describes the time evolution of a probability distribution p_i on a set of discrete states. Mathematically, the rate of change of probability is (2.17)

$$\mathrm{d}_t p_i = \sum_j R_{ij} p_j \ . \tag{A.1}$$

Here, the transition rate matrix elements R_{ij} represent the rate of transitions from state $j \to i$. To ensure conservation of probability, the diagonal elements satisfy $R_{ii} = -\sum_{j \neq i} R_{ij}$.

To propagate the probability distribution forward in time, we use an Euler expansion of the master equation, and approximate the derivative to first-order in the timestep Δt as

$$p_i(t + \Delta t) \approx p_i(t) + \left[\sum_j R_{ij} p_j(t) \right] \Delta t \tag{A.2}$$

which, by iteratively repeating, determines the probability distribution at a later time $\tau = N\Delta t$. Given that the timestep Δt is sufficiently small, this approximation accurately represents the time-evolution of probability in the system.

Given a long-enough time (and assuming time-independent rates R_{ij}) the probability distribution will converge to a steady-state p_i^{ss} which is independent of time (so that $\mathrm{d}_t p_i^{\mathrm{ss}} = 0$). In principle, this distribution can be obtained by iterating (A.2) until the derivative converges to zero, however, there is a simpler way of obtaining the steady-state distribution.

[1] All of the code and data used in this thesis will be made available within reasonable request, and barring any objections from collaborators.

In particular, by viewing the master equation as a matrix equation, the steady-state distribution is calculated by solving for the unique right eigenvector of the rate matrix R_{ji} with 0 eigenvalue [84]

$$\sum_j R_{ji} p_i^{\mathrm{ss}} = 0 \ , \tag{A.3}$$

for all elements i of the probability vector.

A.1.1 Trajectory simulation

In order to generate trajectories of a system governed by master equation dynamics, we make use of a kinetic Monte Carlo method (which is very similar to the well-known Gillespie algorithm [133]). Specifically, for a system currently in discrete state x_i, we assume an exponential dwell-time distribution, and thus draw a random dwell time from the distribution

$$\Delta t = V_{\mathrm{tot}} e^{-V_{\mathrm{tot}} \Delta t} \ , \tag{A.4}$$

where $V_{\mathrm{tot}} = \sum_j V_{ji}$ is the *total rate*, the sum of the rate of all possible transitions out of state x_i.

The next step is to figure out which particular transition occurs next (after a dwell time Δt). To calculate this, we first construct a vector $\tilde{V}$ of the rates in decreasing magnitude—so that $\tilde{V}_i > \tilde{V}_j$ implies that $i > j$—and use this to construct a new vector $\tilde{V}^{\mathrm{cumul}}$ with elements $\tilde{V}_i^{\mathrm{cumul}}$ that are the normalized cumulative sum over all rates greater than (and inclusive of) element $\tilde{V}_i$. Specifically, the ith element of $\tilde{V}^{\mathrm{cumul}}$ is

$$\tilde{V}_i^{\mathrm{cumul}} = \frac{1}{V^{\mathrm{tot}}} \sum_{n=0}^{i} \tilde{V}_n \ . \tag{A.5}$$

We then draw a random variable uniformly in the interval $r \in [0, 1]$ and find the index j for which the inequality

$$\tilde{V}_j^{\mathrm{cumul}} \leq r \leq \tilde{V}_{j+1}^{\mathrm{cumul}} \tag{A.6}$$

is uniquely satisfied. Whichever event corresponds to the jth entry of the vector $\tilde{V}$ is the particular event that occurs after the time interval Δt, and thus we update the system state from $x_i \to x_{i'}$.

This procedure allows us to generate trajectories which are consistent with the master equation given in (A.1).

A.2 Langevin equation

The Langevin equation forms the basis for much of the continuous-space simulations that were presented in this thesis. In all cases considered, a Brownian particle is diffusing in a one-dimensional potential $E(x|\lambda)$, and thus we restrict our attention here to univariate Langevin equations. There are two different classes of Langevin dynamics that we have used in this thesis: underdamped and overdamped.

A.2.1 Underdamped dynamics

For an underdamped system, there are two coupled Langevin equations that describe the time evolution of the stochastic dynamics (2.32)

$$\mathrm{d}x = v\,\mathrm{d}t \tag{A.7a}$$

$$m\mathrm{d}v = f(t)\mathrm{d}t - \frac{1}{\beta D}v\,\mathrm{d}t + \sqrt{\frac{2}{\beta^2 D}}\,\mathrm{d}W(t) \tag{A.7b}$$

where v is the velocity of a diffusing particle, x is its position, $\beta = (k_\mathrm{B}T)^{-1}$ is the inverse temperature, D is the diffusion coefficient, and $W(t)$ is a Wiener process with $\langle W(t)\rangle = 0$ and $\langle W(t)W(t')\rangle = \delta(t-t')$.

To simulate the time evolution of the underdamped particle, we need to discretize the equation of motion. However, there are many different ways to discretize the equations of motion in (A.7), as shown in [233]. Furthermore, a naive implementation of discretization in driven nonequilibrium systems can lead to systematic biases [234]. A systematic study of various different discretization schemes and how they each satisfy different physical *desiderata* is given in Ref. [235], where it was found that the stochastic analog of the popular velocity-Verlet deterministic integrator [133] is given by the following stochastic update sequence[2]

$$v\left(t+\tfrac{1}{4}\Delta t\right) = \sqrt{a}v(t) + \sqrt{\frac{1-a}{\beta m}}\mathcal{N}^{+}(t) \tag{A.8a}$$

$$v\left(t+\tfrac{1}{2}\Delta t\right) = v\left(t+\tfrac{1}{4}\Delta t\right) + \frac{b\Delta t}{2}\frac{f(t)}{m} \tag{A.8b}$$

$$x\left(t+\tfrac{1}{2}\Delta t\right) = x(t) + \frac{b\Delta t}{2}v\left(t+\tfrac{1}{2}\Delta t\right) \tag{A.8c}$$

$$E(x|\lambda_t) \to E(x|\lambda_{t+\Delta t}) \tag{A.8d}$$

$$x\left(t+\Delta t\right) = x\left(t+\tfrac{1}{2}\Delta t\right) + \frac{b\Delta t}{2}v\left(t+\tfrac{1}{2}\Delta t\right) \tag{A.8e}$$

$$v\left(t+\tfrac{3}{4}\Delta t\right) = v\left(t+\tfrac{1}{2}\Delta t\right) + \frac{b\Delta t}{2}\frac{f(t+\Delta t)}{m} \tag{A.8f}$$

$$v\left(t+\Delta t\right) = \sqrt{a}v\left(t+\tfrac{3}{4}\Delta t\right) + \sqrt{\frac{1-a}{\beta m}}\mathcal{N}^{-}(t+\Delta t)\ , \tag{A.8g}$$

where $\mathcal{N}^{\pm}$ is a Gaussian-distributed random variable with zero mean and unit variance, and $f(t)$ is the force experienced by the particle at time t. Here, the coefficient $a \equiv \exp(-\gamma\Delta t)$, where $\gamma \equiv (\beta D)^{-1}$ is the viscous friction coefficient, and thus $a \in [0,1]$ interpolates between the deterministic ($a = 1$) and overdamped ($a = 0$) limits. For $a \approx 1$, the velocities at consecutive timesteps are highly correlated (due to the small impact of the randomization on the velocity updates) and thus are underdamped, while for $a \approx 0$, the velocities at

[2]This integrator is referred to as the 'OVRVO' integrator in Ref. [235], which references the particular Strang splitting of the time-evolution operator, with the legend: 'O'=Ornstein-Uhlenbeck (velocity randomization), 'V'= deterministic velocity update, and 'R'= deterministic position update.

consecutive steps are completely randomized, and the dynamics become overdamped. The coefficient b is an effective timestep-rescaling parameter, which ensures that the relationship between the diffusion coefficient D and the mean-squared displacement is given by the Einstein relation (1.7). It was shown in [235] that

$$b = \sqrt{\frac{2}{\gamma \Delta t} \tanh \frac{\gamma \Delta t}{2}} \ . \tag{A.9}$$

A.2.2 Overdamped dynamics

For an overdamped system, the Langevin equation describing the time evolution of the stochastic dynamics is

$$\mathrm{d}x = -\beta D \partial_x E(x|\lambda)\, \mathrm{d}t + \sqrt{2D}\, \mathrm{d}W(t) \ , \tag{A.10}$$

where $W(t)$ is a Wiener process with $\langle W(t) \rangle = 0$ and $\langle W(t) W(t') \rangle = \delta(t - t')$ (Sec. 2.2.3).

Given an initial state x_t, the update rule (known as the Euler-Maruyama method) we use to generate trajectories of an overdamped system is given by the small-Δt expansion of (A.10) [236]

$$x_{t+\Delta t} \approx x_t - \beta D \partial_x E(x_t|\lambda_t) \Delta t + \sqrt{2D\Delta t}\mathcal{N} \ , \tag{A.11}$$

where x_t and λ_t are the position and control-parameter value at time t, and $\mathcal{N}$ is a Gaussian random variable with zero mean and unit variance.

In the high-friction limit ($a = 1$), the integrator in (A.8) simplifies to the Euler-Maruyama method (A.11) [235].

A.3 Coupled discrete and continuous dynamics

In Ch. 10 we made use of simulations that couple discrete-state dynamics of a master equation with continuous-space dynamics from an overdamped Langevin equation. Specifically, we simulate the joint dynamics of a mechanochemical molecular machine, where a discrete chemical coordinate λ_i imposes a particular continuous energy landscape on a mechanical coordinate x. Here, the dynamics of the mechanical state at a given chemical coordinate are simulated using the overdamped integrator in Sec. A.2.2. However, generating the corresponding chemical trajectories cannot be done in the same manner as A.1.1, as in that section, there is an implicit assumption that throughout the dwell time Δt between discrete transitions, the transition rates remain constant. Here, the coupled mechanochemical dynamics require—through local detailed balance—that the chemical transition rates depend on the current state of the mechanical system.

Instead, we perform an update to the mechanical state of the system $x_{t-\Delta t} \to x_t$ using (A.11) and then calculate the probability that a system in the current mechanical state x_t will experience a transition over the next time interval Δt, which here is the timestep used for the mechanical dynamics (A.11). Specifically, the probability that a chemical transition

occurs within the next Δt is given by the cumulative probability of dwell times $\Delta t_{\text{dwell}} \in [0, \Delta t]$ (assuming an exponential dwell-time distribution),

$$p_{\text{exit}}(x_t, \Delta t) = \int_0^{\Delta t} W^{\text{tot}}(x_t) e^{-W^{\text{tot}}(x_t)\Delta t} \mathrm{d}\Delta t \tag{A.12a}$$

$$= 1 - W^{\text{tot}}(x_t) e^{-W^{\text{tot}}(x_t)\Delta t} \Delta t \ , \tag{A.12b}$$

where $W^{\text{tot}}(x) = \sum_j W_{ji}(x)$ is the sum over all position-dependent exit rates $W_{ji}(x_t)$ from chemical state λ_i for a system currently in mechanical state x_t. The exit probability gets small as $\Delta t \to 0$, thus after each mechanical state update, we draw a uniform random number r between zero and one; if $r < p_{\text{exit}}(x_t, \Delta t)$, then a transition occurs, and the particular transition that occurs is determined in the same manner as Sec. A.1.1, thus reproducing dynamics that are consistent with the coupled discrete-continuous mechanochemical system.

Appendix B

DNA hairpins

B.1 Folding forces

In Sec. 4.1 we compared the difference between the unfolding/refolding forces $f_{\text{naive}} - f_{\text{designed}}$ of a naive and designed protocol. Given that the unfolding/refolding force was used as a proxy for the distance from equilibrium of the hairpin along a particular protocol, the unfolding/refolding force difference informs us how much more out of equilibrium the naive protocol is relative to the designed protocol of the same duration. We also pointed out that, by comparing two force-separation curves of the same duration from the same molecule, we cancel out any experimental artifacts due to inter-molecule variation or instrument calibration errors. However, there are also other comparisons that can be made between folding forces that can be used to arrive at similar conclusions.

For instance, the difference between the folding force along an unfolding protocol and a refolding protocol of the same class (designed or naive)—which we call the *force hysteresis*—also cancels out such experimental artifacts, and also acts as a measure of the distance from equilibrium. However, instead of directly comparing the unfolding or refolding forces between a designed and naive protocol, this compares the overall distance from equilibrium—including both unfolding and refolding directions—and compares this measure between designed and naive protocols. Unlike the difference considered in Sec. 4.1, however, the long-duration limit will not be zero, but will approach a finite value representing the difference between the local maximum and minimum forces in the vicinity of a force-rip event (which here appears to be just above 2 pN).

Figure B.1 shows a comparison of the force hysteresis for naive and designed protocols in the fast-relaxing hairpin for a variety of protocol durations. In Fig. B.1a, force hysteresis histograms show that as the protocol duration decreases (and thus the average protocol velocity increases), the naive cycles have increased force hysteresis, relative to designed cycles. Furthermore, Fig. B.1b shows that, over all durations, the designed protocols lead to a lower average force hysteresis, and asymptote to a nonzero value of ≈ 2 pN in the long-duration limit.

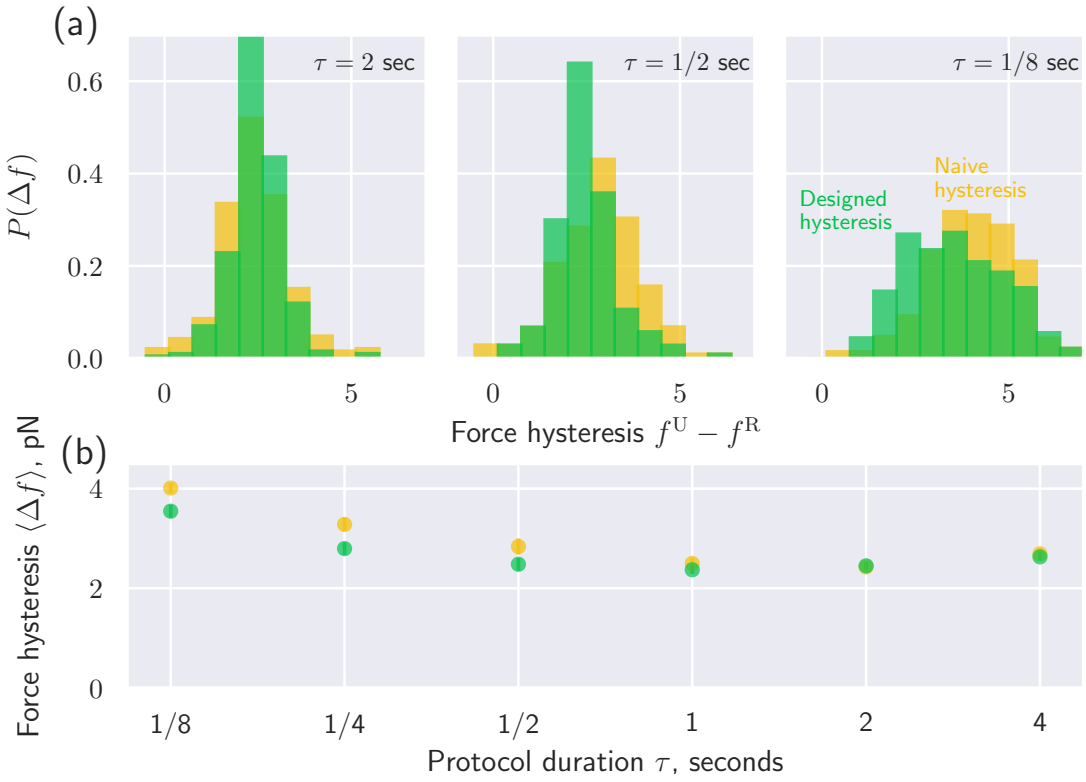

Figure B.1: **Naive/designed force hysteresis in a DNA hairpin.** (a) Histograms of force hysteresis for naive (yellow) and designed (green) protocols. (b) Average force hysteresis for naive (yellow) and designed (green) protocols, across a range of protocol durations, asymptoting in the long-duration limit to a nonzero value slightly greater than ~ 2 pN.

B.2 Alternative excess work measures

In Sec. 4.2, we used the cycle work $W_{\mathrm{cycle}} \equiv W^{\mathrm{U}} + W^{\mathrm{R}}$ as a measure of (twice) the excess work. Here, the contribution to the work of a given unfolding protocol due to the free energy difference cancels out when added to the work of a refolding protocol, and thus the cycle work is equal to twice the excess work $W_{\mathrm{cycle}} = 2W_{\mathrm{ex}}$ (assuming that the excess work along an unfolding or refolding protocol are the same).

However, there are alternative ways of combining work measurements to cancel out the free energy difference that could also be used, such as the work difference between a naive and a designed protocol, each in a particular direction (unfolding or refolding). Here, the free energy is canceled, and the result is equal to the difference in excess work between a naive and designed protocol

$$W_{\mathrm{diff}} = W_{\mathrm{naive}} - W_{\mathrm{designed}} \tag{B.1a}$$
$$= W_{\mathrm{ex,naive}} - W_{\mathrm{ex,designed}} + \Delta F - \Delta F \tag{B.1b}$$
$$= W_{\mathrm{ex,naive}} - W_{\mathrm{ex,designed}} \ , \tag{B.1c}$$

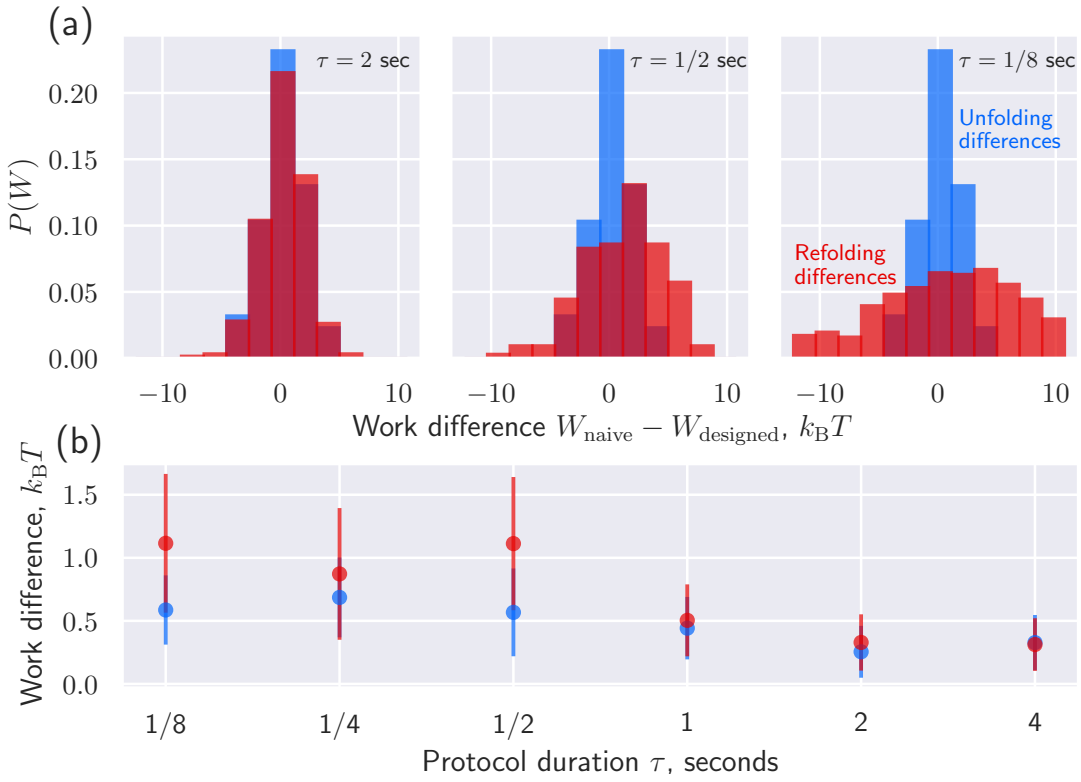

Figure B.2: **Work differences in fast-relaxing hairpin.** (a) The distributions of work differences $W_{\text{naive}} - W_{\text{designed}}$ for unfolding (blue) and refolding (red) protocols, for protocols ranging from slow (left) to fast (right). (b) Mean work difference $\langle W_{\text{naive}} - W_{\text{designed}} \rangle$ during unfolding (blue) and refolding (red) protocols as a function of protocol duration. Error bars indicate twice the standard error of the mean.

and thus directly addresses how much a designed protocol of a particular duration outperforms its naive counterpart.

Figure B.2 shows the histograms of work differences and the mean work difference over several protocol durations. In Fig. B.2a, as the protocol duration decreases, histograms for both the unfolding and refolding differences shift to higher values and the variance increases. Additionally, the refolding differences are more significantly impacted by the change in protocol duration than the unfolding differences, reflecting the different physical processes involved in unfolding or refolding the hairpin. Figure B.2b shows the trend of mean differences as a function of protocol duration. Here, values greater than zero indicate that the designed protocols are saving work relative to the naive counterpart. We see—over all durations investigated—that there is a positive average work difference, which gets larger as the protocol duration gets smaller (and thus the average velocity increases). Furthermore, the average work differences for unfolding and refolding protocols are statistically indistinguishable over the durations that we have considered, even though the variances of histograms shown in Fig. B.2a are notably different. These results provide further indication that the designed protocols systematically reduce excess work when compared to their naive counterparts.

Appendix C

Stochastic Control in Microscopic Nonequilibrium Systems

C.1 Generalization of lower dissipation bound

Written explicitly for a single control parameter, our lower bound on excess work (6.29) is

$$\langle W_{\rm ex}\rangle_\Omega \geq 2\left\langle \zeta(\langle\lambda\rangle_{\Omega_t})\langle\delta\dot{\lambda}^2\rangle_{\Omega_t}\right\rangle_\Omega^{1/2}\int_0^\tau \sqrt{\zeta(\langle\lambda\rangle_{\Omega_t})\langle\dot{\lambda}\rangle^2_{\Omega_t}}\mathrm{d}t\ . \tag{C.1}$$

Machta's lower bound on the entropy production of a stochastically driven system is [180]

$$\langle\Sigma\rangle_\Omega \geq 2\int_0^\tau \sqrt{\mathcal{I}(\langle\lambda\rangle_{\Omega_t})\langle\dot{\lambda}\rangle^2_{\Omega_t}}\mathrm{d}t\ , \tag{C.2}$$

where $\mathcal{I}(\langle\boldsymbol{\lambda}\rangle) = \beta^2\langle\delta f^2\rangle_{\langle\boldsymbol{\lambda}\rangle}$ is the Fisher information matrix and $\langle\Sigma\rangle_\Omega$ is the entropy production averaged over the protocol ensemble Ω. When control-parameter manipulation is the only source of entropy production, and the integral relaxation time $\tau^{\rm R}$ is constant along the protocol, we rewrite (C.2) as an excess work

$$\begin{aligned}\langle W_{\rm ex}\rangle_\Omega &= k_{\rm B}T\langle\Sigma\rangle_\Omega &\text{(C.3a)}\\ &\geq 2k_{\rm B}T\int_0^\tau \sqrt{\mathcal{I}(\langle\lambda\rangle_{\Omega_t})\langle\dot{\lambda}\rangle^2_{\Omega_t}}\mathrm{d}t &\text{(C.3b)}\\ &= 2\sqrt{\frac{k_{\rm B}T}{\tau^{\rm R}}}\int_0^\tau \sqrt{\zeta(\langle\lambda\rangle_{\Omega_t})\langle\dot{\lambda}\rangle^2_{\Omega_t}}\mathrm{d}t\ . &\text{(C.3c)}\end{aligned}$$

The two lower bounds (C.1) and (C.3c) are thus equivalent when

$$\tau^{\rm R} = \left(\beta\left\langle\zeta(\langle\lambda\rangle_{\Omega_t})\langle\delta\dot{\lambda}^2\rangle_{\Omega_t}\right\rangle_\Omega\right)^{-1}\ . \tag{C.4}$$

To understand when this equality is achieved, consider Machta's total entropy production (of system and control parameter combined)

$$\Sigma_{\text{sys}^+} = \frac{1}{2}\mathcal{I}(\langle\lambda\rangle_{\Omega_t})\Delta\lambda^2 \tag{C.5a}$$

$$= \frac{1}{2}\mathcal{I}(\langle\lambda\rangle_{\Omega_t})(\tau^{\text{R}})^2\left(\frac{\Delta\lambda}{\tau^{\text{R}}}\right)^2 , \tag{C.5b}$$

due to a fluctuation in the protocol (Eq. 9 from [180]). Rewriting this as an excess work

$$W_{\text{ex}}^{\text{stoch}} = \frac{1}{2}\zeta(\langle\lambda\rangle_{\Omega_t})\left(\frac{\Delta\lambda}{\tau^{\text{R}}}\right)^2\tau^{\text{R}} , \tag{C.6a}$$

where we used the decomposition of the generalized friction in (2.66), $\zeta(\lambda) = k_{\text{B}}T\tau^{\text{R}}\mathcal{I}(\lambda)$, and denoted by $W_{\text{ex}}^{\text{stoch}}$ the contribution to the total excess work due to the stochastic fluctuations of the protocol away from the average path. For protocols satisfying the locally deterministic limit discussed in Sec. 6.2,

$$\frac{\Delta\lambda}{\tau^{\text{R}}} \approx \dot{\lambda} . \tag{C.7}$$

I.e., over time scales comparable to the integral relaxation time τ^{R}, the control-parameter velocity is constant. This makes the instantaneous contribution to the excess work due to fluctuations

$$W_{\text{ex}}^{\text{stoch}} \approx \frac{1}{2}\zeta(\langle\lambda\rangle_{\Omega_t})\dot{\lambda}^2\tau^{\text{R}} = \frac{1}{2}\zeta(\langle\lambda\rangle_{\Omega_t})\delta\dot{\lambda}^2\tau^{\text{R}} , \tag{C.8}$$

where the final equality expresses that this is the excess work due to a fluctuation away from the average protocol, so $\dot{\lambda} \to \dot{\lambda} - \langle\dot{\lambda}\rangle_{\Omega_t} = \delta\dot{\lambda}$. In a reference frame comoving with the average protocol, if the distribution of control-parameter velocities is stationary then the instantaneous probability of observing a trajectory with a particular control-parameter fluctuation away from the average is

$$P(\delta\dot{\lambda}) \propto \exp\left(-\frac{\beta}{2}\zeta(\langle\lambda\rangle_{\Omega_t})\delta\dot{\lambda}^2\tau^{\text{R}}\right) , \tag{C.9}$$

and thus the squared fluctuation at each point along the protocol, averaged over the instantaneous ensemble of control-parameter velocities, is

$$\langle\delta\dot{\lambda}^2\rangle_{\Omega_t} = (\beta\zeta(\langle\lambda\rangle_{\Omega_t})\tau^{\text{R}})^{-1} . \tag{C.10}$$

Rearranging and averaging gives

$$\tau^{\text{R}} = \left(\beta\left\langle\zeta(\langle\lambda\rangle_{\Omega_t})\langle\delta\dot{\lambda}^2\rangle_{\Omega_t}\right\rangle_{\Omega}\right)^{-1} , \tag{C.11}$$

thereby reducing our general lower bound (C.1) to Machta's bound (C.2) [180].

C.2 Disagreement between theoretical predictions and numerical results

Numerical results and theoretical predictions in the bottom row ($\beta E^\ddagger = 4$) of Fig. 6.3 disagree because in any given protocol ensemble, the average excess work is dominated by the most rapid sampled protocols. For these fastest protocols, the (nonequilibrium) system position distribution $p(x,t)$ significantly lags the harmonic-trap minimum. As a result, the system experiences a force dominated by the harmonic trap, with little influence from the underlying periodic potential, so the excess work is well-approximated by a system driven solely by a harmonic trap, as confirmed in Fig. C.1.

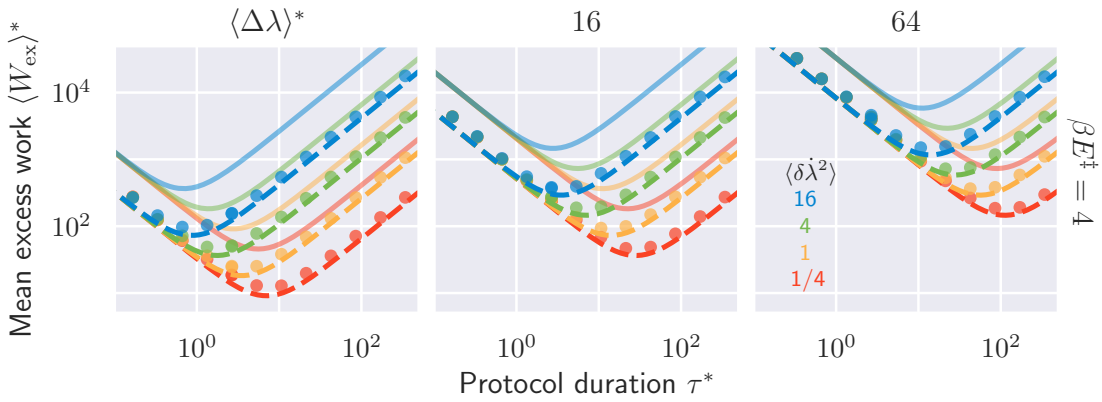

Figure C.1: **Harmonic approximation for excess work agrees with numerical simulations of high-barrier periodic-potential ensemble simulations.** For high-barrier ($\beta E^\ddagger = 4$) simulations of the periodic-potential ensemble (dots), the excess work is well approximated by a system driven by a translating harmonic potential (dashed lines), in contrast to the linear-response theory predictions for a system driven by a harmonic trap translating over an underlying periodic potential (solid lines).

C.3 Equivalence of ensembles

According to the theory presented in Chapter 6, the excess work is a function only of the control-parameter velocity's mean and variance across the protocol ensemble. The primary constraint imposed on the ensembles is that for each protocol $\Lambda \in \Omega$ the excess work can be accurately approximated by (2.69a) [109]. I.e., details of the ensemble, such as boundary conditions, do not affect the excess work.

For instance, the ensemble of stochastic protocols considered in Sec. (6.4.2) has a Brownian Bridge boundary condition on the protocols: each protocol starts and finishes at the same initial and final control-parameter values λ_i and λ_f, respectively, but has a variable duration [237]. We alternatively consider the ensemble in which each protocol starts at the same control-parameter value λ_i and has the same duration, but has a variable final position λ_f (in Sec. 6.4.1). Given that both cases have the same control-parameter velocity mean and

variance, the theoretically predicted excess work is equal, regardless of the substantially different mathematical procedures necessary to find exact solutions. This logic also encompasses the zero-barrier periodic-potential ensemble, where the generalized friction is the same as the stochastic-protocol ensembles, and thus the excess work values are predicted to be equal. However, it is not possible to draw this equivalence with the nonzero-barrier periodic-potential ensemble, because it has a different average protocol.

Figure C.2 shows sample trajectories from each of these ensembles with equal theoretically predicted excess works (Fig. C.2a) and numerically demonstrates the equivalent excess works in the appropriate limit (Fig. C.2b). In particular, for large protocol distance (Fig. C.2b, right panel), over all protocol durations the control-parameter dynamics satisfy the linear-response (2.69a) and locally deterministic (6.4) approximations, and hence the different ensembles produce identical mean work that also matches the theoretical approximation derived in the main text (6.27).

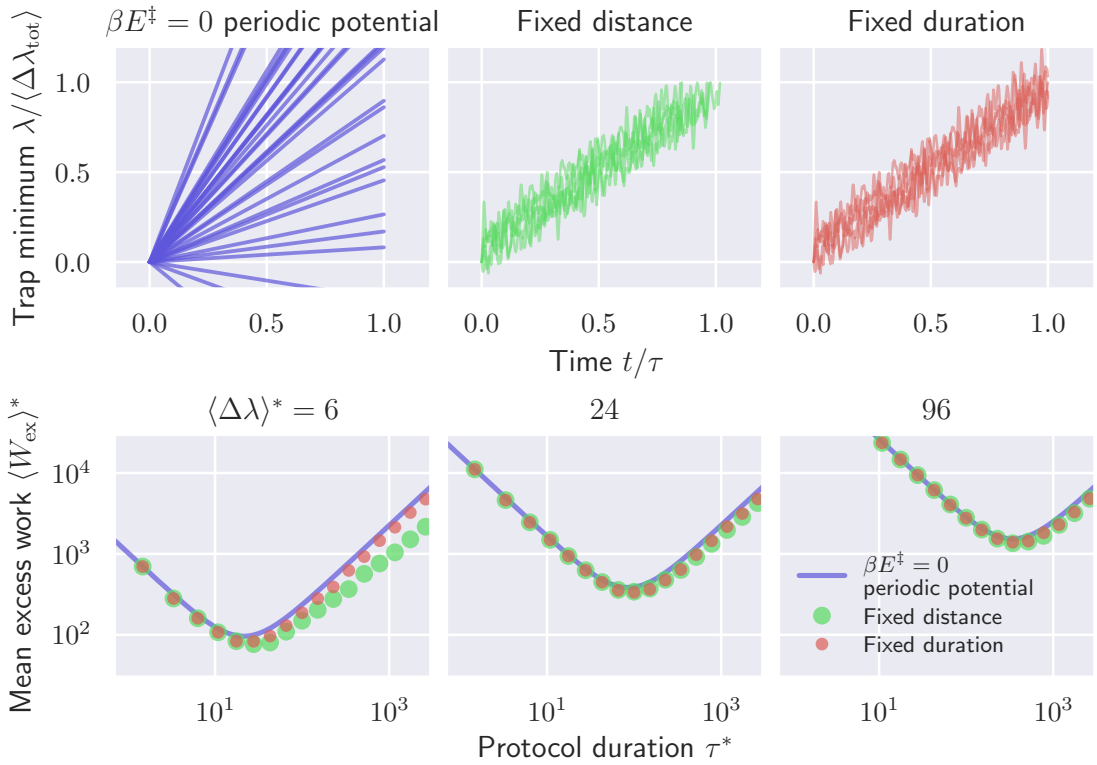

Figure C.2: **For near-deterministic protocol ensembles, the excess work depends only on the average protocol and the control-parameter velocity variance.** (a) Sample trajectories from the zero-barrier ($\beta E^\ddagger = 0$) periodic-potential ensemble (left, blue), and from a stochastic protocol ensemble with boundary condition of either fixed protocol distance (middle, red) or fixed protocol duration (right, green) with $D_\lambda^* \equiv D_\lambda/D = 25$. In all cases $\langle\delta\dot{\lambda}^2\rangle_\Omega^* = 1$. (b) Average excess work for each ensemble, as a function of protocol duration, with colors matching those in top row. Right: excess work is indeed identical across ensembles in the limits when the linear-response (2.69a) and locally deterministic (6.4) approximations hold.

Interestingly, the statistical moments that appear in (6.27) are those of the control-parameter velocity, not its position. As a result, the precision of the protocol distance, defined as the inverse variance of the final position, can vary significantly depending on the choice of ensemble, while maintaining the same energetic cost. For instance, the periodic-potential protocol ensemble has a precision that decreases secularly with protocol duration, while the stochastically driven protocol ensemble maintains a bound precision regardless of the protocol duration.

Appendix D

Optimal discrete control: minimizing dissipation in discretely driven nonequilibrium systems

D.1 Expansion of the relative entropy

The relative entropy (Kullback-Leibler divergence) between two continuous probability distributions $p(x)$ and $q(x)$ is defined as [94]

$$D[q(x)||p(x)] \equiv \int \ln\left[\frac{p(x)}{q(x)}\right] p(x)\mathrm{d}x \ . \tag{D.1}$$

In the context of the present work, the equilibrium distribution $\pi(x|\boldsymbol{\lambda}_i)$ is parameterized by the control parameter $\boldsymbol{\lambda}$. The integrand of the relative entropy for two consecutive equilibrium distributions at $\boldsymbol{\lambda}_i$ and $\boldsymbol{\lambda}_{i+1}$ is

$$g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}) = \pi(x|\boldsymbol{\lambda}_i) \ln \frac{\pi(x|\boldsymbol{\lambda}_i)}{\pi(x|\boldsymbol{\lambda}_{i+1})}. \tag{D.2}$$

For small changes $\Delta\boldsymbol{\lambda}_{i,i+1} \equiv \boldsymbol{\lambda}_{i+1} - \boldsymbol{\lambda}_i$ in the control parameter, we Taylor expand Eq. (D.2) about $\boldsymbol{\lambda}_i$,

$$\begin{aligned} g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}) = g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_i) &+ \left[\partial_{\lambda^j_{i+1}} g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1})\right]_{\lambda^j_i} \Delta\lambda^j_{i,i+1} \\ &+ \frac{1}{2}\left[\partial_{\lambda^j_{i+1}} \partial_{\lambda^k_{i+1}} g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1})\right]_{\lambda^j_i, \lambda^k_i} \Delta\lambda^j_{i,i+1}\Delta\lambda^k_{i,i+1} + \mathcal{O}(\Delta\boldsymbol{\lambda}^3) \ , \end{aligned} \tag{D.3}$$

where $\partial_{\lambda^m_{i+1}} g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1}) \equiv \frac{\partial}{\partial\lambda^m_{i+1}} g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1})$ is the partial derivative of $g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i+1})$ with respect to the mth component of the control parameter $\boldsymbol{\lambda}_{i+1}$, $[\cdots]_{\lambda^m_i}$ indicates that the argument is evaluated at $\lambda^m_{i+1} = \lambda^m_i$, and we have made use of the Einstein summation notation, where repeated indices are summed over.

The first term in (D.3) is

$$g(x, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_i) = \pi(x|\boldsymbol{\lambda}_i) \ln \frac{\pi(x|\boldsymbol{\lambda}_i)}{\pi(x|\boldsymbol{\lambda}_i)} = \pi(x|\boldsymbol{\lambda}_i) \ln 1 = 0 \ . \tag{D.4}$$

The derivative on the RHS of (D.3), evaluated at $\boldsymbol{\lambda}_{i+1} = \boldsymbol{\lambda}_i$ is

$$\partial_{\lambda^j_{i+1}} \left\{ \pi(x|\boldsymbol{\lambda}_i) \ln \frac{\pi(x|\boldsymbol{\lambda}_i)}{\pi(x|\boldsymbol{\lambda}_{i+1})} \right\}_{\lambda^j_i} = -\frac{\partial \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^j_{i+1}}\bigg|_{\lambda^j_i} \ , \tag{D.5}$$

and the second-derivative term in (D.3) is

$$\begin{aligned} \partial_{\lambda^j_{i+1}} \partial_{\lambda^k_{i+1}} \left\{ \pi(x|\boldsymbol{\lambda}_i) \ln \frac{\pi(x|\boldsymbol{\lambda}_i)}{\pi(x|\boldsymbol{\lambda}_{i+1})} \right\}_{\lambda^j_i, \lambda^k_i} &= \frac{1}{\pi(x|\boldsymbol{\lambda}_i)} \left[\frac{\partial \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^j_{i+1}} \frac{\partial \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^k_{i+1}} \right]_{\lambda^j_i, \lambda^k_i} \\ &\quad - \frac{\partial^2 \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^j_{i+1} \partial \lambda^k_{i+1}}\bigg|_{\lambda^j_i, \lambda^k_i} \ . \end{aligned} \tag{D.6}$$

Equation (D.5) can be simplified by noting that the equilibrium probability distribution is normalized, $\int \pi(x, \boldsymbol{\lambda}) \mathrm{d}x = 1$, and partial differentiation commutes with integration, so substituting (D.5) into the relative entropy expression (D.1), gives

$$\frac{\partial}{\partial \lambda^j_{i+1}} \int \pi(x|\boldsymbol{\lambda}_{i+1}) \, \mathrm{d}x = \frac{\partial}{\partial \lambda^j_{i+1}} 1 = 0 \ , \tag{D.7}$$

so this term does not contribute to the overall relative entropy. This results from the relative entropy being a convex function with a minimum at $\Delta\boldsymbol{\lambda} = 0$. In analogy with (D.7), the second term on the RHS of (D.6) also vanishes,

$$\frac{\partial^2}{\partial \lambda^k_{i+1} \partial \lambda^j_{i+1}} \int \pi(x|\boldsymbol{\lambda}_{i+1}) \, \mathrm{d}x = \frac{\partial^2}{\partial \lambda^k_{i+1} \partial \lambda^j_{i+1}} 1 = 0 \ . \tag{D.8}$$

Combining (D.7) and (D.8) with (D.3), the relative entropy for a control-parameter step $\Delta\boldsymbol{\lambda}_{i,i+1}$ is

$$D\left[\pi(x|\boldsymbol{\lambda}_i)||\pi(x|\boldsymbol{\lambda}_{i+1})\right] = \frac{1}{2} \Delta\lambda^j_{i,i+1} \Delta\lambda^k_{i,i+1} \int \frac{1}{\pi(x|\boldsymbol{\lambda}_i)} \left[\frac{\partial \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^j_{i+1}} \frac{\partial \pi(x|\boldsymbol{\lambda}_{i+1})}{\partial \lambda^k_{i+1}} \right]_{\boldsymbol{\lambda}_i} \mathrm{d}x + \mathcal{O}(\Delta\boldsymbol{\lambda}^3) \ , \tag{D.9}$$

where the integral is the Fisher information matrix $\mathcal{I}_{jk}(\boldsymbol{\lambda}_i)$ at control-parameter value $\boldsymbol{\lambda}_i$ [94]. For sufficiently small steps, the $\mathcal{O}(\Delta\boldsymbol{\lambda}^3)$ term is negligible, so for a discrete control protocol Λ, consisting of N steps, the excess work in the infinite-time limit (7.9b) is

$$\langle \beta W_{\mathrm{ex}} \rangle_\Lambda = \frac{1}{2} \sum_{i=0}^{N-1} \mathcal{I}_{jk}(\boldsymbol{\lambda}_i) \Delta\lambda^j_{i,i+1} \Delta\lambda^k_{i,i+1} \ . \tag{D.10}$$

For a physical system in contact with a thermal reservoir, the equilibrium distribution is (1.1). Within the linear-response regime the energy can be expanded about $\boldsymbol{\lambda}_0$ (7.12b),

$$E(x,\boldsymbol{\lambda}) \approx E(x,\boldsymbol{\lambda}_0) - f_j|_{\boldsymbol{\lambda}_0}(\lambda^j - \lambda_0^j) + \mathcal{O}(\Delta\boldsymbol{\lambda}^2) \tag{D.11}$$

for the conjugate force $f_j|_{\boldsymbol{\lambda}_0} \equiv -\partial E/\partial\lambda^j|_{\boldsymbol{\lambda}_0}$ (evaluated at $\boldsymbol{\lambda}_0$). Derivatives of the equilibrium distribution are

$$\partial_{\lambda^j}\pi(x|\boldsymbol{\lambda}) = \beta\left(f_j|_{\boldsymbol{\lambda}} + \frac{\partial F(\boldsymbol{\lambda})}{\partial\lambda^j}\right)\pi(x|\boldsymbol{\lambda}) \ . \tag{D.12}$$

From the thermodynamic definition of the free energy,

$$F(\boldsymbol{\lambda}) = \langle E\rangle_{\boldsymbol{\lambda}} - TS \tag{D.13}$$

$$= -\lambda^j\langle f_j\rangle_{\boldsymbol{\lambda}} - TS \ , \tag{D.14}$$

so partial derivatives of the free energy in (D.12) are $\partial_{\lambda^j}F(\boldsymbol{\lambda}) = -\langle f_j\rangle_{\boldsymbol{\lambda}}$. Therefore, for an equilibrium distribution (1.1), the Fisher information is

$$\mathcal{I}_{jk}(\boldsymbol{\lambda}_i) = \int \frac{1}{\pi(x|\boldsymbol{\lambda}_i)}\left[\partial_{\lambda^j}\pi(x|\boldsymbol{\lambda})\partial_{\lambda^k}\pi(x|\boldsymbol{\lambda})\right]_{\boldsymbol{\lambda}_i} \tag{D.15}$$

$$= \beta^2\int \left(f_j|_{\boldsymbol{\lambda}_i} - \langle f_j\rangle_{\boldsymbol{\lambda}_i}\right)\left(f_k|_{\boldsymbol{\lambda}_k} - \langle f_k\rangle_{\boldsymbol{\lambda}_i}\right)\pi(x|\boldsymbol{\lambda}_i)\,\mathrm{d}x \tag{D.16}$$

$$= \beta^2\langle\delta f_j\delta f_k\rangle_{\boldsymbol{\lambda}_i} \ , \tag{D.17}$$

where $\delta f_j \equiv f_j|_{\boldsymbol{\lambda}_i} - \langle f_j\rangle_{\boldsymbol{\lambda}_i}$. Substituting (D.17) into (D.10) gives

$$\langle\beta W_{\mathrm{ex}}\rangle_\Lambda = \frac{1}{2}\beta^2\sum_{i=0}^{N-1}\langle\delta f_j\delta f_k\rangle_{\boldsymbol{\lambda}_i}\Delta\lambda^j_{i,i+1}\Delta\lambda^k_{i,i+1} \ . \tag{D.18}$$

This final equation is equivalent to the infinite-time protocol work (7.11) in Sec. 7.3.

D.2 Harmonic trap: exact result

For a system in contact with a thermal bath, subjected to a harmonic confining potential,

$$E_{\mathrm{trap}}(x,\lambda) = \frac{1}{2}k_{\mathrm{trap}}\,(x-\lambda)^2 \ , \tag{D.19}$$

and initially at equilibrium, the initial distribution over microstates is

$$\pi(x|\lambda_0) = \sqrt{\frac{\beta k_{\mathrm{trap}}}{2\pi}}\,e^{-\frac{1}{2}\beta k_{\mathrm{trap}}(x-\lambda_0)^2} \ . \tag{D.20}$$

The protocol work can be calculated exactly when this system is subjected to a discrete control protocol, which takes the trap minimum through a sequence of positions $\lambda_0, \lambda_1, \cdots, \lambda_N$ (with fixed spring constant).

The first step $\lambda_0 \to \lambda_1$ requires average work

$$\langle W \rangle_{\lambda_0 \to \lambda_1} = \int \left[E(x, \lambda_1) - E(x, \lambda_0) \right] \pi(x|\lambda_0) \mathrm{d}x \tag{D.21}$$

$$= \frac{1}{2} k_{\mathrm{trap}} \Delta\lambda_{0,1}^2 \ . \tag{D.22}$$

After the control-parameter change, the system is in a nonequilibrium distribution given by the solution to the one-dimensional Fokker-Planck equation

$$\partial_t p(x, t|\lambda_1, \lambda_0) = -\beta D k_{\mathrm{trap}} \partial_x \left[(x - \lambda_1) p(x, t|\lambda_1) \right] + \frac{1}{2} D \partial_{xx}^2 p(x, t|\lambda_1) \ , \tag{D.23}$$

subject to the initial condition $p(x, t = 0|\lambda_1, \lambda_0) = \pi(x|\lambda_0)$. ($D$ is the system diffusion coefficient.) The exact solution is known [72]: after a time Δt_1 spent at λ_1, the time-dependent probability distribution is

$$p(x, t_1|\lambda_1, \lambda_0) = \sqrt{\frac{\beta k_{\mathrm{trap}}}{2\pi}} \exp\left\{ -\frac{1}{2} \beta k_{\mathrm{trap}} \left(x - \lambda_1 + \Delta\lambda_{0,1} e^{-\beta D k_{\mathrm{trap}} \Delta t_1} \right)^2 \right\} \ , \tag{D.24}$$

a Gaussian distribution with time-dependent mean $\lambda_1 - \Delta\lambda_{0,1} e^{-\beta D k_{\mathrm{trap}} \Delta t_1}$, which approaches λ_1 in the infinite-time ($\Delta t_1 \to \infty$) limit.

If, after time Δt_1, the second control parameter step $\lambda_1 \to \lambda_2$ takes place, the resulting average work is

$$\langle W \rangle_{\lambda_1 \to \lambda_2} = \int \left[E(x, \lambda_2) - E(x, \lambda_1) \right] p(x, t_1|\lambda_1, \lambda_0) \mathrm{d}x \tag{D.25}$$

$$= \frac{1}{2} k_{\mathrm{trap}} \Delta\lambda_{1,2}^2 + k_{\mathrm{trap}} \Delta\lambda_{1,2} \Delta\lambda_{0,1} e^{-\beta D k_{\mathrm{trap}} \Delta t_1} \ . \tag{D.26}$$

Again, after the control-parameter change $\lambda_1 \to \lambda_2$, the system is out of equilibrium with probability distribution solving the Fokker-Planck equation (D.23), subject to the initial condition $p(x, t = 0|\lambda_2, \lambda_1, \lambda_0, \Delta t_1) = p(x, \Delta t_1|\lambda_1, \lambda_0)$. This leads to the time-dependent system distribution at λ_2 after a time Δt_2,

$$p(x, \Delta t_2|\lambda_2, \lambda_1, \lambda_0, \Delta t_1) = \sqrt{\frac{\beta k_{\mathrm{trap}}}{2\pi}} \exp\left\{ -\frac{1}{2} \beta k_{\mathrm{trap}} \left(x - \lambda_2 + \xi_1 e^{-\beta D k_{\mathrm{trap}} \Delta t_2} \right)^2 \right\} \tag{D.27}$$

for $\xi_1 \equiv \Delta\lambda_{1,2} + \Delta\lambda_{0,1} e^{-\beta D k_{\mathrm{trap}} \Delta t_1}$. For all subsequent steps ($\lambda_{i-1} \to \lambda_i$) the probability distribution has the same form as (D.27)

$$p(x, \Delta t_i|\lambda_i, \lambda_{i-1}, \cdots, \lambda_0, \Delta t_{i-1}, \cdots, \Delta t_1) \tag{D.28}$$

$$= \sqrt{\frac{\beta k_{\mathrm{trap}}}{2\pi}} \exp\left\{ -\frac{1}{2} \beta k_{\mathrm{trap}} \left(x - \lambda_i + \xi_{i-1} e^{-\beta D k_{\mathrm{trap}} \Delta t_i} \right) \right\}$$

for

$$\xi_{i-1} \equiv \sum_{n=0}^{i-1} \Delta\lambda_{n,n+1} \exp\left(-\beta D k_{\mathrm{t}} \sum_{m=n+1}^{i-1} \Delta t_m \right) \ . \tag{D.29}$$

For the time-dependent distribution in (D.28), the average work during the step $\lambda_i \to \lambda_{i+1}$ is

$$\langle W \rangle_{\lambda_i \to \lambda_{i+1}} = \frac{1}{2} k_{\text{trap}} \Delta\lambda_{i,i+1}^2 + k_{\text{trap}} \Delta\lambda_{i,i+1} \xi_{i-1} e^{-\beta D k_{\text{t}} \Delta t_i} \ . \tag{D.30}$$

From (D.30), the total work during an arbitrary discrete protocol for a harmonic potential is

$$\langle W \rangle_{\Lambda} = \sum_{i=0}^{N-1} k_{\text{trap}} \Delta\lambda_{i,i+1}^2 \left[\frac{1}{2} + \frac{\xi_{i-1}}{\Delta\lambda_{i,i+1}} e^{-\beta D k_{\text{trap}} \Delta t_i} \right] \ . \tag{D.31}$$

Appendix E

On dissipation bounds

E.1 Generalized friction for Gamma-distributed dwell times

For control-parameter jump dynamics, the replacement of an exponential dwell-time distribution with a different functional form will change the amount of excess work. As a model for such a dwell-time distribution, we use the Gamma distribution [193]

$$p_{ji}^{\text{dwell}}(\Delta t) = \frac{(V_{ji})^{\alpha}}{\Gamma(\alpha)}(\Delta t)^{\alpha-1}e^{-V_{ji}\Delta t} \ , \tag{E.1}$$

where α is a shape parameter, and $\Gamma(\alpha)$ is the Gamma function evaluated at α. This distribution represents a two-parameter generalization of the standard exponential distribution, reducing to it for $\alpha = 1$.

Using the Gamma distribution in place of the exponential dwell-time distribution in the nonequilibrium excess work during a discrete jump (8.4) gives

$$\langle \beta W_{\text{ex}}^{\text{neq}} \rangle_{ji|\ell}^{\Omega} \approx \beta^2 \Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}\frac{(V_{ji})^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}\langle \delta f_k(0)\delta f_{k'}(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}(\Delta t_i)^{\alpha-1}e^{-V_{ji}\Delta t_i}\mathrm{d}\Delta t_i \ . \tag{E.2}$$

By again assuming that the control-parameter dynamics are slow relative to the conjugate-force relaxation and approximating the exponential term $\exp(-V_{ji}\Delta t_i) \approx 1$, we can rewrite (E.2) as

$$\langle \beta W_{\text{ex}}^{\text{neq}} \rangle_{ji|\ell}^{\Omega} \approx \beta^2 \Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}\frac{(V_{ji})^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}\langle \delta f_k(0)\delta f_{k'}(\Delta t_i)\rangle_{\boldsymbol{\lambda}_i}(\Delta t_i)^{\alpha-1}\mathrm{d}\Delta t_i \tag{E.3a}$$

$$= \beta\Delta\lambda_{ji}^{k}\Delta\lambda_{i\ell}^{k'}(V_{ji})^{\alpha}\zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}_i) \tag{E.3b}$$

where we define the *non-exponential* generalization of the friction tensor

$$\zeta_{kk'}^{(\alpha)}(\boldsymbol{\lambda}) \equiv \frac{\beta}{\Gamma(\alpha)}\int_0^{\infty} t^{\alpha-1}\langle \delta f_k(0)\delta f_{k'}(t)\rangle_{\boldsymbol{\lambda}}\mathrm{d}t \tag{E.4}$$

which quantifies the resistance that the system will put up to changes in the control parameter for non-exponential jump dynamics, where the dwell-time distribution is given by (E.1).

The non-exponential friction tensor reduces to the standard friction tensor presented in [109] for $\alpha = 1$, where the dwell-time distribution is exponential.

E.2 Generalized friction for Gamma-distributed dwell times: harmonic-trap

For a the quadratic potential given in (2.73), the linearity of the conjugate forces allows us to express the force autocovariance function as

$$\langle \delta f(0) \delta f(t) \rangle_\lambda = k^2 \langle \delta x(0) \delta x(t) \rangle_\lambda \tag{E.5a}$$

$$= \frac{k}{\beta} e^{-\frac{k}{\gamma} t} \,, \tag{E.5b}$$

where in (E.5b) we use the position autocovariance for an Ornstein-Uhlenbeck process [72]. We can then calculate $\zeta^{(\alpha)}$ explicitly,

$$\zeta^{(\alpha)} = \frac{k}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-\frac{k}{\gamma} t} \mathrm{d}t \tag{E.6a}$$

$$= k_{\mathrm{trap}} \left(\frac{\gamma}{k_{\mathrm{trap}}} \right)^\alpha \tag{E.6b}$$

$$= k_{\mathrm{trap}}^{1-\alpha} \gamma^\alpha \,, \tag{E.6c}$$

which for $\alpha = 1$ reduces to the usual form of $\zeta = \gamma$.

E.3 Average step number for uniform jump rates

Here, we derive the expected step number $\langle N_\Lambda \rangle_\Omega$ for the biased random walk model system investigated in the main text (Sec. 8.4). In particular, for an ensemble of stochastic control protocols traveling between the initial position λ_0 and λ_N through a series of discrete jumps of size $\Delta\lambda$, which occur with forward (+) and reverse (−) rates $V_\pm$, we are interested in the average total number of steps taken during the protocol. We can also define the forward and reverse jump probabilities

$$p_+ = \frac{V_+}{V_+ + V_-} \tag{E.7a}$$

$$p_- = \frac{V_-}{V_+ + V_-} \tag{E.7b}$$

which are, respectively, the probability that the next step in the Markov chain will be in the positive (+) or negative (−) direction. The boundary values of the protocol are separated by a distance $\lambda_N - \lambda_0 = N_{\mathrm{Fw}} \Delta\lambda$, and each individual realization of a stochastic protocol Λ takes a total number of steps $N_\Lambda \geq N_{\mathrm{Fw}}$.

Mathematically, this problem can be first simplified by calculating the first-passage time of the random walker to state N_{Fw}. We define the random variable $X_i = \pm 1$ as a counter,

with the index i representing the number of steps taken up to that point. The position of the random walker at step n equals the nth partial sum of the process,

$$S_n \equiv \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n \ , \tag{E.8}$$

bound from above by $S_n < N_{\mathrm{Fw}}$.

To calculate the average step number $\langle N_\Lambda \rangle_\Omega$, we first calculate the expected number of steps required to take a single forward step $\langle \min_\Lambda [N|S_N = 1] \rangle_\Omega$, and use the recursive property of first-passage times on Markov chains along with the equivalence of each step in the simple mode to express the average step number as

$$\langle N_\Lambda \rangle_\Omega = N_{\mathrm{Fw}} \left\langle \min_\Lambda [N|S_N = 1] \right\rangle_\Omega \ . \tag{E.9}$$

We will denote the event that a first passage occurs after n steps as F_n, which has the moment-generating function

$$\phi(z) = \langle z^N \rangle = \sum_{n=1}^{\infty} z^n p(F_n) \ . \tag{E.10}$$

Now, given that the jumps are Markovian, we can use the fact that for two independent events ϵ_1 and ϵ_2 we can write the expectation of z^N in (E.10) as

$$\langle z^N \rangle = \langle z^N | X_1 = 1 \rangle p(X_1 = 1) + \langle z^N | X_1 = -1 \rangle p(X_1 = -1) \ . \tag{E.11}$$

The first RHS term is simple, as $\langle z^N | X_1 = 1 \rangle = z$ and thus $\langle z^N | X_1 = 1 \rangle P(X_1 = 1) = z P_+$. To evaluate the second RHS term, we introduce two new random variables $\tilde{N}_1, \tilde{N}_2$, which represent respectively the number of steps to go from $X_1 = -1$ to $X_{\tilde{N}_1} = 0$, and $X_{\tilde{N}_1} = 0$ to $X_{\tilde{N}_2} = 1$. In terms of these new random variables, we can write

$$\langle z^N | X_1 = -1 \rangle = \langle z^{1+\tilde{N}_1+\tilde{N}_2} | X_1 = -1 \rangle \tag{E.12a}$$

$$= z \langle z^N \rangle \langle z^N \rangle \tag{E.12b}$$

$$= z \phi^2(z) \ , \tag{E.12c}$$

where (E.12b) follows from the mutual independence of $\tilde{N}_1$ and $\tilde{N}_2$, and (E.12c) follows from (E.10).

Putting (E.12c) back into (E.11) gives the quadratic equation $z p_- \phi(z)^2 - \phi(z) + p_+ z = 0$, which can be solved to give

$$\phi(z) = \frac{1 \pm \sqrt{1 - 4 p_+ p_- z^2}}{2 p_- z} \ . \tag{E.13}$$

To then find the expectation $\langle z^N \rangle$, we take the first derivative of (E.13) and evaluate it at $z = 1$:

$$\partial_z \phi(z)|_{z=1} = \left[\sum_{n=0}^{\infty} n z^{n-1} p(F_n) \right]_{z=1} \tag{E.14a}$$

$$= \sum_{n=0}^{\infty} n p(F_n) \tag{E.14b}$$

$$= \langle N_\Lambda \rangle \ . \tag{E.14c}$$

For the moment-generating function in (E.13) we get

$$\partial_z \phi(z)|_{z=1} = \frac{1 - \sqrt{1 - 4p_+ p_- z^2}}{2p_- z^2 \sqrt{1 - 4p_+ p_- z^2}} \Big|_{z=1} \tag{E.15a}$$

$$\langle N \rangle = \frac{p_+ + p_- - \sqrt{(p_+ - p_-)^2}}{2p_- \sqrt{(p_+ - p_-)^2}} \tag{E.15b}$$

$$= \frac{p_+ + p_- - |p_+ - p_-|}{2p_- |p_+ - p_-|} \ , \tag{E.15c}$$

where we assume $p_+ \geq p_-$ and use the facts that $p_+ + p_- = 1$ and $1 = (p_+ + p_-)^2 = p_+ + p_- + 2p_+ p_-$, which implies that $1 - 4p_+ p_- = (p_+ - p_-)^2$. If $p_+ = p_-$ the solution diverges ($\langle N \rangle \to \infty$), as expected, and for $p_+ > p_-$ we get

$$\langle N \rangle = \frac{1}{p_+ - p_-} = \frac{V_+ + V_-}{V_+ - V_-} \ . \tag{E.16}$$

Finally, making use of the recursive property of first-passage times, we find the protocol-ensemble averaged total step number for a fixed-distance protocol with N_{Fw} net forward steps as

$$\langle N_\Lambda \rangle_\Omega = N_{\mathrm{Fw}} \frac{V_+ + V_-}{V_+ - V_-} \ . \tag{E.17}$$

Appendix F

Free energy transduction within autonomous systems

F.1 Detailed derivation of transduced additional free energy rate

At steady state, the rate of change of internal energy $\langle E \rangle_{XY}$ is zero, thus

$$\begin{aligned} 0 &= \mathrm{d}_t \langle E \rangle_{XY} \qquad \text{(F.1a)} \\ &= \sum_{x,y,y'} \beta(\epsilon_{xy} - \epsilon_{xy'}) R_{yy'} x p_{xy'} + \sum_{x,x',y} \beta(\epsilon_{xy} - \epsilon_{x'y}) R_y x x' p_{x'y} \ . \end{aligned}$$

Substituting this into the entropy production due to X dynamics [Eq. (9.5)], we separate the energetic contribution:

$$\dot{\Sigma}^X = \beta \sum_{x,y,y'} R_{yy'} x p_{xy'} \left(\epsilon_{xy} - \epsilon_{xy'}\right) - \beta \, \mathrm{d}_t \langle E \rangle_{XY} + \sum_{x,x',y} R_y x x' p_{x'y} \ln \frac{p_{x'|y}}{p_{x|y}} \ . \qquad \text{(F.2)}$$

We identify the first RHS term as the power defined in Eq. (9.9), while the second and third RHS terms can only be combined to a derivative of free energy if we include $\dot{I}^Y$, capturing the change in mutual information due to the Y dynamics:

$$\begin{aligned} \beta \, \mathrm{d}_t \langle E \rangle_{XY} - \sum_{x,x',y} R_y x x' p_{x'y} \ln \frac{p_{x'|y}}{p_{x|y}} &= \beta \, \mathrm{d}_t \langle E \rangle_{XY} - \sum_{x,x',y,y'} R_{yy'} x x' p_{x'y'} \ln \frac{p_{x'|y'}}{p_{x|y}} \qquad \text{(F.3a)} \\ &\quad + \sum_{x,y,y'} R_{yy'} x p_{xy'} \ln \frac{p_{x|y'}}{p_{x|y}} \\ &= \beta \, \mathrm{d}_t F^{\mathrm{neq}}_{X|Y} - \dot{I}^Y \ . \qquad \text{(F.3b)} \end{aligned}$$

Here, we used the definition [Eq. (9.7b)] of the information rate due to Y dynamics, and used the conditional nonequililbirium free energy from Eq. (9.24).

Consequently, we find:

$$\dot{\Sigma}^X = \langle \beta \mathcal{P} \rangle_{Y \to X} - \beta \, \mathrm{d}_t F^{\mathrm{neq}}_{X|Y} + \dot{I}^Y \geq 0 \; . \tag{F.4}$$

The transduced additional free energy rate is then defined using the conditional *equilibrium* free energy $F_{X|Y}$ (the average over Y of $F_{X|y}$) instead of its nonequilibrium counterpart:

$$\beta \dot{F}^{\mathrm{add}}_{Y \to X} = \langle \beta \mathcal{P} \rangle_{Y \to X} - \beta \, \mathrm{d}_t F_{X|Y} + \dot{I}^Y \; . \tag{F.5}$$

We can now more fully appreciate the similarity between transduced additional free energy rate and entropy production rate: For processes starting and ending in equilibrium, integrating the entropy production rate and the excess power give the same result. For steady-state systems, both measures agree because both the equilibrium and nonequilibrium free energy are unchanging. Furthermore, using the equilibrium free energy in Eq. (F.5) results in an expression for the transduced additional free energy rate which is independent of X dynamics, while the same is not true if the nonequilibrium free energy $F^{\mathrm{neq}}_{X|Y}$ [Eq. (9.24)] is used, as its rate of change depends on the X dynamics through the conditional entropy $S_{X|Y}$.

F.2 At steady state, excess power equals heat flow

Here we derive–for detailed-balanced dynamics–the equality of the entropy production in the reservoir (heat flow) due to the dynamics of subsystem X, and the excess power done on subsystem X by subsystem Y. We use Eq. (F.1) to rewrite the excess power [Eq. (9.9)] as

$$\langle \beta \mathcal{P}_{\mathrm{ex}} \rangle_{Y \to X} = \sum_{x,y,y'} \beta(\epsilon_{xy} - \epsilon_{xy'}) R_{yy'} x p_{xy'} \tag{F.6a}$$

$$= - \sum_{x,x',y} \beta(\epsilon_{xy} - \epsilon_{x'y}) R_y x x' p_{x'y} \tag{F.6b}$$

$$= \sum_{x,x',y} R_y x x' p_{x'y} \ln \frac{\pi_{x|y}}{\pi_{x'|y}} \tag{F.6c}$$

$$= \sum_{x,x',y} R_y x x' p_{x'y} \ln \frac{R_y x x'}{R_y x' x} \tag{F.6d}$$

$$= -\langle \beta \dot{\mathcal{Q}} \rangle^X \; . \tag{F.6e}$$

In Eq. (F.6c) we use Eq. (9.10) and the fact that the rate of change of conditional free energy is zero, as the conditional distributions are unchanging. In Eq. (F.6d) we relate the log ratio of conditional equilibria to the microscopic rates. Finally, in Eq. (F.6e) we substitute the definition of heat flow from Ref. [195].

Appendix G

Hidden excess power and autonomous Maxwell demons in nonequilibrium systems

G.1 Expansion of the TSS work

While the expressions for the TSS excess work in (10.16b) are exact, it is convenient to examine limiting cases so as to approximate the excess work in terms of equilibrium averages. For instance, the expansion of the relative entropy in App. D.1 shows that, for small $\Delta\boldsymbol{\lambda}$, the conjugate-force variance can be used to approximate the excess work associated with a discrete transition, when the discrete transitions are independent of the mechanical state $\boldsymbol{x}$. Here, we proceed similarly, Taylor expanding the integrand of (10.16b) in $\Delta\boldsymbol{\lambda}$. To start we define

$$g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j) \equiv \sqrt{\pi_i \pi_j} \ln \frac{\pi_i}{\pi_j} \ , \tag{G.1}$$

the integrand of (10.16b). For small steps $\Delta\boldsymbol{\lambda}_{ji} = \boldsymbol{\lambda}_j - \boldsymbol{\lambda}_i$, we Taylor expand (G.1) in $\boldsymbol{\lambda}_j$ about $\boldsymbol{\lambda}_i$,

$$\begin{aligned} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j) \approx g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_i) + \left[\partial_{\lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)\right]_{\lambda_i^k} \Delta\lambda_{ji}^k + \frac{1}{2}\left[\partial^2_{\lambda_j^\ell \lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i \boldsymbol{\lambda}_j)\right]_{\lambda_i^k \lambda_i^\ell} \Delta\lambda_{ji}^k \Delta\lambda_{ji}^\ell \\ + \frac{1}{3!}\left[\partial^3_{\lambda_j^m \lambda_j^\ell \lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)\right]_{\lambda_i^k \lambda_i^\ell \lambda_i^m} \Delta\lambda_{ji}^k \Delta\lambda_{ji}^\ell \Delta\lambda_{ji}^m + \mathcal{O}(\Delta\boldsymbol{\lambda}_{ji}^4) \ , \end{aligned} \tag{G.2}$$

where $\partial_{\lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j) \equiv \frac{\partial}{\partial_{\lambda_j^k}} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)$ is the partial derivative of $g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)$ with respect to the kth component of the $\boldsymbol{\lambda}$ vector, and $[\cdots]_{\lambda_i^k}$ indicates that the argument is evaluated at $\lambda_j^k = \lambda_i^k$. We have also made use of Einstein summation notation, where repeated indices are implicitly summed over.

The first RHS term is

$$g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_i) = \pi_i(\boldsymbol{x}) \ln \frac{\pi_i(\boldsymbol{x})}{\pi_i(\boldsymbol{x})} = \pi_i \ln 1 = 0 \ . \tag{G.3}$$

The RHS first-derivative term in (G.2) is

$$\partial_{\lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j) = \left(\partial_{\lambda_j^k} \sqrt{\pi_i \pi_j}\right) \ln \frac{\pi_i}{\pi_j} + \sqrt{\pi_i \pi_j} \left(\partial_{\lambda_j^k} \ln \frac{\pi_i}{\pi_j}\right) \tag{G.4a}$$

$$= \frac{1}{2} \left(\partial_{\lambda_j^k} \pi_j\right) \sqrt{\frac{\pi_i}{\pi_j}} \ln \frac{\pi_i}{\pi_j} - \sqrt{\pi_i \pi_j} \frac{1}{\pi_j} \left(\partial_{\lambda_j^k} \pi_j\right) \tag{G.4b}$$

which when evaluated at $\lambda_j^k = \lambda_i^k$ gives

$$\left[\partial_{\lambda_j^k} g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)\right]_{\lambda_i^k} = \frac{1}{2} \left[\partial_{\lambda_j^k} \pi_j\right]_{\lambda_i^k} \ln 1 - \left[\partial_{\lambda_j^k} \pi_j\right]_{\lambda_i^k} \tag{G.5a}$$

$$= - \left[\partial_{\lambda_j^k} \pi_j\right]_{\lambda_i^k} \tag{G.5b}$$

.

Substituting (G.4) into the RHS second-derivative term of (G.2) gives

$$\partial_{\lambda_j^k} \partial_{\lambda_j^\ell} \left[\sqrt{\pi_i \pi_j} \ln \frac{\pi_i}{\pi_j}\right] = \partial_{\lambda_j^\ell} \left[\frac{1}{2} \left(\partial_{\lambda_j^k} \pi_j\right) \sqrt{\frac{\pi_i}{\pi_j}} \ln \frac{\pi_i}{\pi_j} - \sqrt{\frac{\pi_i}{\pi_j}} \left(\partial_{\lambda_j^k} \pi_j\right)\right] \ . \tag{G.6}$$

The second RHS term,

$$\partial_{\lambda_j^\ell} \left[\sqrt{\frac{\pi_i}{\pi_j}} \left(\partial_{\lambda_j^k} \pi_j\right)\right] = \left(\partial_{\lambda_j^k} \sqrt{\frac{\pi_i}{\pi_j}}\right) \left(\partial_{\lambda_j^k} \pi_j\right) + \sqrt{\frac{\pi_i}{\pi_j}} \left(\partial^2_{\lambda_j^\ell, \lambda_j^k} \pi_j\right) \tag{G.7a}$$

$$= -\frac{1}{2} \sqrt{\frac{\pi_i}{\pi_j^3}} \left(\partial_{\lambda_j^\ell} \pi_j\right) \left(\partial_{\lambda_j^k} \pi_j\right) + \sqrt{\frac{\pi_i}{\pi_j}} \left(\partial^2_{\lambda_j^\ell \lambda_j^k} \pi_j\right) \ , \tag{G.7b}$$

evaluated at $\lambda_j^k = \lambda_i^k$ and $\lambda_j^\ell = \lambda_i^\ell$ gives

$$\partial_{\lambda_j^\ell} \left[\sqrt{\frac{\pi_i}{\pi_j}} \left(\partial_{\lambda_j^k} \pi_j\right)\right] = -\frac{1}{2\pi_i} \left[\partial_{\lambda_j^\ell} \pi_j\right]_{\lambda_i^\ell} \left[\partial_{\lambda_j^k} \pi_j\right]_{\lambda_i^k} + \left[\partial^2_{\lambda_j^\ell \lambda_j^k} \pi_j\right]_{\lambda_i^k \lambda_i^\ell} \ . \tag{G.8}$$

The first RHS term of (G.6),

$$\partial_{\lambda_j^\ell} \left[\frac{1}{2} \left(\partial_{\lambda_j^k} \pi_j\right) \sqrt{\frac{\pi_i}{\pi_j}} \ln \frac{\pi_i}{\pi_j}\right] = \frac{1}{2} \sqrt{\frac{\pi_i}{\pi_j}} \ln \left(\frac{\pi_i}{\pi_j}\right) \left(\partial^2_{\lambda_j^\ell \lambda_j^k} \pi_j\right) - \frac{1}{2} \left(\partial_{\lambda_j^\ell} \pi_j\right) \left(\partial_{\lambda_j^k} \pi_j\right) \sqrt{\frac{\pi_i}{\pi_j^3}}$$

$$- \frac{1}{4} \left(\partial_{\lambda_j^\ell} \pi_j\right) \left(\partial_{\lambda_j^k} \pi_j\right) \sqrt{\frac{\pi_i}{\pi_j^3}} \ln \frac{\pi_i}{\pi_j} \tag{G.9}$$

evaluated at $\lambda_j^k = \lambda_i^k$ and $\lambda_j^\ell = \lambda_i^\ell$ gives

$$\partial_{\lambda_j^\ell}\left[\frac{1}{2}\left(\partial_{\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j}}\ln\frac{\pi_i}{\pi_j}\right]_{\lambda_i^k\lambda_i^\ell} = \frac{1}{2}\left[\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right]_{\lambda_i^\ell\lambda_i^k}\ln 1 - \frac{1}{2\pi_i}\left[\partial_{\lambda_j^k}\pi_i\right]_{\lambda_i^k}\left[\partial_{\lambda_j^\ell}\pi_j\right]_{\lambda_i^\ell}$$
$$- \frac{1}{4\pi_i}\left[\partial_{\lambda_j^k}\pi_i\right]_{\lambda_i^k}\left[\partial_{\lambda_j^\ell}\pi_j\right]_{\lambda_i^\ell}\ln 1 \tag{G.10a}$$
$$= -\frac{1}{2\pi_i}\left[\left(\partial_{\lambda_j^k}\pi_j\right)\left(\partial_{\lambda_j^\ell}\pi_j\right)\right]_{\lambda_i^k\lambda_i^\ell} . \tag{G.10b}$$

Substituting (G.8) and (G.10) into (G.6) gives

$$\partial^2_{\lambda_j^\ell\lambda_j^k}\left[\sqrt{\pi_i\pi_j}\ln\frac{\pi_i}{\pi_j}\right]_{\lambda_i^\ell\lambda_i^k} = -\left[\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right]_{\lambda_i^\ell\lambda_j^k} . \tag{G.11}$$

Substituting (G.11), (G.5), and (G.3) into (G.2) gives the $\mathcal{O}(\Delta\boldsymbol{\lambda}^2)$ approximation of $g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j)$:

$$g(\boldsymbol{x}, \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j) \approx -\left[\partial_{\lambda_j^k}\pi_j\right]_{\lambda_i^k}\Delta\lambda_{ji}^k - \frac{1}{2}\left[\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right]_{\lambda_i^\ell\lambda_i^k}\Delta\lambda_{ji}^k\Delta\lambda_{ji}^\ell + \mathcal{O}(\Delta\boldsymbol{\lambda}^3) . \tag{G.12}$$

Both derivatives of π_j vanish upon integration over $\boldsymbol{x}$:

$$\int \partial_{\lambda_j^k}\pi_j \,\mathrm{d}\boldsymbol{x} = \partial_{\lambda_j^k}\int \pi_j \,\mathrm{d}\boldsymbol{x} = 0 \tag{G.13a}$$
$$\int \partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j \,\mathrm{d}\boldsymbol{x} = \partial_{\lambda_j^\ell}\partial_{\lambda_j^k}\int \pi_j \,\mathrm{d}\boldsymbol{x} = 0 , \tag{G.13b}$$

indicating that $\langle\beta W_{\mathrm{ex}}^{\mathrm{TSS}}\rangle_{ij}$ is third order in $\Delta\boldsymbol{\lambda}$: in contrast to the discrete-control TSS excess work in [77], the second-order term vanishes.

Substituting (G.7) and (G.9) into (G.6) and differentiating gives the third-derivative term in (G.2):

$$\partial^3_{\lambda_j^m\lambda_j^\ell\lambda_j^k}\left[\sqrt{\pi_i\pi_j}\ln\frac{\pi_i}{\pi_j}\right] = \partial_{\lambda_j^m}\left[\left(\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j}}\phi_{ji} - \frac{1}{4}\left(\partial_{\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j^3}}\ln\frac{\pi_i}{\pi_j}\right] . \tag{G.14}$$

We have defined

$$\phi_{ji} \equiv \frac{1}{2}\ln\frac{\pi_i}{\pi_j} - 1 , \tag{G.15}$$

such that

$$\partial_{\lambda_j^m}\phi_{ji} = -\frac{1}{\pi_j}\left(\partial_{\lambda_j^m}\pi_j\right) , \tag{G.16}$$

and at $\boldsymbol{\lambda}_j = \boldsymbol{\lambda}_i$, $\phi_{ji} = -1$ and $\partial_{\lambda_j^m}\phi_{ji} = -\frac{1}{2\pi_i}[\partial_{\lambda_j^m}\pi_j]_{\lambda_i^m}$.

Thus the first RHS term of (G.14) is

$$\partial_{\lambda_j^m}\left[\left(\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j}}\phi_{ji}\right] = \left(\partial^3_{\lambda_j^m\lambda_j^\ell\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j}}\phi_{ji} + \left(\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right)\left(\partial_{\lambda_j^m}\phi_{ji}\right)\sqrt{\frac{\pi_i}{\pi_j}} - \frac{1}{2}\left(\partial^2_{\lambda_j^\ell\lambda_j^k}\pi_j\right)\left(\partial_{\lambda_j^m}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j^3}}\phi_{ji} \tag{G.17a}$$

$$= -\left[\partial^3_{\lambda_j^m\lambda_j^\ell\lambda_j^k}\pi_j\right]_{\lambda_i^m\lambda_i^\ell\lambda_i^k}\ , \quad \boldsymbol{\lambda}_j = \boldsymbol{\lambda}_i\ . \tag{G.17b}$$

The second RHS term of (G.14) is

$$\partial_{\lambda_j^m}\left[\left(\partial_{\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j^3}}\ln\frac{\pi_i}{\pi_j}\right]$$

$$= \left(\partial^2_{\lambda_j^m\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j^3}}\ln\frac{\pi_i}{\pi_j} + \left(\partial_{\lambda_j^\ell\pi_j}\pi_j\right)\left(\partial^2_{\lambda_j^m\lambda_j^k}\pi_j\right)\sqrt{\frac{\pi_i}{\pi_j^3}}\ln\frac{\pi_i}{\pi_j} \tag{G.18a}$$

$$- \frac{1}{3}\left(\partial_{\lambda_j^m}\pi_j\right)\left(\partial_{\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_i\right)\sqrt{\frac{\pi_i}{\pi_j^4}}\ln\frac{\pi_i}{\pi_j} - \left(\partial_{\lambda_j^m}\pi_j\right)\left(\partial_{\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_i\right)\sqrt{\frac{\pi_i}{\pi_j^5}}$$

$$= -\frac{1}{\pi_i^2}\left[\partial_{\lambda_j^m}\pi_j\right]_{\lambda_j^m}\left[\partial_{\lambda_j^\ell}\pi_j\right]_{\lambda_j^\ell}\left[\partial_{\lambda_j^k}\pi_j\right]_{\lambda_j^k}\ , \quad \boldsymbol{\lambda}_j = \boldsymbol{\lambda}_i\ . \tag{G.18b}$$

Substituting (G.18b) and (G.17b) into (G.14) gives

$$\partial^3_{\lambda_j^m\lambda_j^\ell\lambda_j^k}\left[\sqrt{\pi_i\pi_j}\ln\frac{\pi_i}{\pi_j}\right] = -\left[\partial^3_{\lambda_j^m\lambda_j^\ell\lambda_j^k}\pi_j\right]_{\lambda_i^m\lambda_i^\ell\lambda_i^k} + \frac{1}{4\pi_i^2}\left[\partial_{\lambda_j^m}\pi_j\right]_{\lambda_j^m}\left[\partial_{\lambda_j^\ell}\pi_j\right]_{\lambda_j^\ell}\left[\partial_{\lambda_j^k}\pi_j\right]_{\lambda_j^k}\ . \tag{G.19}$$

The first RHS term vanishes upon integration for the same reasons as (G.13). Thus, the leading-order approximation of (G.1) in the small-$\Delta\boldsymbol{\lambda}$ limit is

$$\langle\beta W_{\text{ex}}^{\text{TSS}}\rangle_{ji} \approx \frac{\Delta\lambda_{ji}^m\Delta\lambda_{ji}^\ell\Delta\lambda_{ji}^k}{4!}\int\frac{1}{\pi_i^2}\left[\left(\partial_{\lambda_j^m}\pi_j\right)\left(\partial_{\lambda_j^\ell}\pi_j\right)\left(\partial_{\lambda_j^k}\pi_j\right)\right]_{\boldsymbol{\lambda}_i}\mathrm{d}x\ . \tag{G.20}$$

Using $\partial_{\lambda_j^k}\ln\pi_j = \frac{1}{\pi_j}\partial_{\lambda_j^k}\pi_j$ (and evaluating the resulting expression at $\boldsymbol{\lambda}_j = \boldsymbol{\lambda}_i$) simplifies the integral to the expectation of three derivatives of the log-probability,

$$\langle\beta W_{\text{ex}}^{\text{TSS}}\rangle_{ji} \approx \frac{\Delta\lambda_{ji}^m\Delta\lambda_{ji}^\ell\Delta\lambda_{ji}^k}{4!}\left\langle\left(\partial_{\lambda_j^m}\ln\pi_j\right)\left(\partial_{\lambda_j^\ell}\ln\pi_j\right)\left(\partial_{\lambda_j^k}\ln\pi_j\right)\right\rangle_{\boldsymbol{\lambda}_i} \tag{G.21a}$$

$$= \frac{\Delta\lambda_{ji}^m\Delta\lambda_{ji}^\ell\Delta\lambda_{ji}^k}{4!}\mathcal{S}_{m\ell k}(\boldsymbol{\lambda}_i)\ , \tag{G.21b}$$

where third-rank tensor $\mathcal{S}_{m\ell k}(\boldsymbol{\lambda})$ is the leading-order non-Gaussian approximation of the log-probability in the limit of small $\Delta\boldsymbol{\lambda}$ [238]. $\mathcal{S}_{m\ell k}$ is related to the *flexion tensor* used in the analysis of astrophysical image data. The name 'flexion tensor' derives from the original use of flexion as a measure of third-order distortions in astrophysical images due to weak gravitational lensing [239, 240].

For a physical system in contact with a heat reservoir, the derivative of the log-probability is

$$\partial_{\lambda^k} \ln \pi = \beta \left(f_k|_{\boldsymbol{\lambda}} + \partial_{\lambda^k} F(\boldsymbol{\lambda}) \right) \tag{G.22a}$$

$$= \delta f_k|_{\boldsymbol{\lambda}} \tag{G.22b}$$

where $f_k|_{\boldsymbol{\lambda}} \equiv -\partial_{\lambda^k} E(x|\boldsymbol{\lambda})|_{\boldsymbol{\lambda}}$ is the generalized force conjugate to control parameter λ^k at control-parameter vector $\boldsymbol{\lambda}$, and $F(\boldsymbol{\lambda}) = \langle E \rangle_{\boldsymbol{\lambda}} - TS = -\lambda^k \langle f_k \rangle_{\boldsymbol{\lambda}} - TS$ is the equilibrium free energy at $\boldsymbol{\lambda}$.

Substituting (G.22) into the TSS excess work (G.21b) gives

$$\langle \beta W_{\text{ex}}^{\text{TSS}} \rangle_{ji} \approx \frac{1}{4!} \Delta\lambda_{ji}^m \Delta\lambda_{ji}^\ell \Delta\lambda_{ji}^k \langle \delta f_m \delta f_\ell \delta f_k \rangle_{\boldsymbol{\lambda}_i} \ , \tag{G.23}$$

for third centered moment $\langle \delta f_m \delta f_\ell \delta f_k \rangle_{\boldsymbol{\lambda}_i}$ of the generalized forces at control-parameter vector $\boldsymbol{\lambda}_i$.

This analysis shows two primary features: in the TSS limit the excess work due to autonomous systems discretely transitioning between states can be calculated exactly through (10.17), and in the small-$\Delta\boldsymbol{\lambda}$ limit, the leading-order contribution to the TSS excess work (G.21b) is $\mathcal{O}(\Delta\boldsymbol{\lambda}^3)$. Furthermore, for physical systems in contact with thermal reservoirs, the third-rank flexion tensor can be expressed as a matrix of third centered moments of the conjugate forces (G.23).

G.2 Nonequilibrium excess work in autonomous systems

Here we consider the additional excess work in a system driven out of equilibrium by biased chemical dynamics. In particular, we write the total excess work associated with a particular chemical transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ as the sum of two components

$$\langle \beta W_{\text{ex}} \rangle_{ji} = \langle \beta W_{\text{ex}}^{\text{TSS}} \rangle_{ji} + \langle \beta W_{\text{ex}}^{\text{neq}} \rangle_{ji} \ , \tag{G.24}$$

where $\langle \beta W_{\text{ex}}^{\text{TSS}} \rangle_{ji}$ is the excess work in the timescale-separated limit (Appendix G.1), and $\langle \beta W_{\text{ex}}^{\text{neq}} \rangle_{ji}$ is the *nonequilibrium excess work*, the additional excess work required of the chemical dynamics due to the system being out of equilibrium.

In order to give a general form for the nonequilibrium excess work, we appeal to linear-response theory. Specifically, we evaluate the integral expression

$$\langle W \rangle_{ji} = \int \left[E(\boldsymbol{x}|\boldsymbol{\lambda}_j) - E(\boldsymbol{x}|\boldsymbol{\lambda}_i) \right] p_{\text{neq}}(\boldsymbol{x}, t) p_{ji}^{\text{dwell}}(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t \ , \tag{G.25}$$

the total work done on the mechanical degrees of freedom by the chemical dynamics for a transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, averaged over mechanical states $\boldsymbol{x}$ and times t. $p_{\text{neq}}(\boldsymbol{x}, t)$ is the nonequilibrium distribution over mechanical states $\boldsymbol{x}$ at time t, and $p_{ji}^{\text{dwell}}(\boldsymbol{x}, t)$ is the distribution of dwell times for the $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ transitions in mechanical state $\boldsymbol{x}$. We assume that at the

microstate level, the chemical jump dynamics are Markovian with dwell-time distribution

$$p_{ji}^{\text{dwell}}(\boldsymbol{x},t) = R_{ji}^{\boldsymbol{x}} e^{-R_{ji}^{\boldsymbol{x}} t} \ , \tag{G.26}$$

with rates $R_{ji}^{\boldsymbol{x}}$ given in (10.3).

We make a weak-perturbation approximation, Taylor expanding the energy landscape $E(\boldsymbol{x}|\boldsymbol{\lambda}_j)$ around $E(\boldsymbol{x}|\boldsymbol{\lambda}_i)$:

$$E(\boldsymbol{x}|\boldsymbol{\lambda}_j) \approx E(\boldsymbol{x}|\boldsymbol{\lambda}_i) + \left[\nabla_{\boldsymbol{\lambda}} \cdot E(\boldsymbol{x}|\boldsymbol{\lambda})\right]_{\boldsymbol{\lambda}_i} (\boldsymbol{\lambda}_j - \boldsymbol{\lambda}_i) \tag{G.27a}$$

$$= E(\boldsymbol{x}|\boldsymbol{\lambda}_i) - f_k|_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^k \ . \tag{G.27b}$$

This simplifies the rates (10.3) to

$$R_{ji}^{\boldsymbol{x}} = \Gamma_{ji}\, e^{-\frac{1}{2}\beta[E(\boldsymbol{x}|\boldsymbol{\lambda}_j) - E(\boldsymbol{x}|\boldsymbol{\lambda}_i)]} \tag{G.28a}$$

$$\approx \Gamma_{ji}\, e^{\frac{1}{2}\beta f_k|_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^k} \tag{G.28b}$$

$$\approx \Gamma_{ji}\, \left(1 + \tfrac{1}{2}\beta f_k|_{\boldsymbol{\lambda}_i} \Delta\lambda_{ji}^k\right) \ , \tag{G.28c}$$

where $\Gamma_{ji} \equiv \Gamma \exp\left(-\frac{1}{2}\beta\Delta\mu_{ji}\right)$. Substituting (G.26), (G.28b,c), and (G.27) simplifies the excess work (G.25) to

$$\langle W \rangle_{ji} \approx -\Gamma_{ji}\Delta\lambda_{ji}^k \int f_k|_{\boldsymbol{\lambda}_i}(\boldsymbol{x}) p_{\text{neq}}(\boldsymbol{x},t) e^{-\frac{1}{2}\beta\Delta E_{ji}} e^{-\Gamma_{ji}\left(1+\frac{1}{2}\beta\Delta E_{ji}\right)t}\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t \tag{G.29a}$$

$$\approx -\Gamma_{ji}\Delta\lambda_{ji}^k \int_0^\infty e^{-\Gamma_{ji}t} \left\{ \sum_{n=0}^\infty \frac{(-1)^n}{n!} (\Gamma_{ji}t)^n \int f_k|_{\boldsymbol{\lambda}_i}(\boldsymbol{x}) \left(\tfrac{1}{2}\beta\Delta E_{ji}\right)^n p_{\text{neq}}(\boldsymbol{x},t) e^{\frac{1}{2}\beta\Delta E_{ji}}\, \mathrm{d}\boldsymbol{x} \right\} \mathrm{d}t \ , \tag{G.29b}$$

where $\Delta E_{ji} \equiv f_k|_{\boldsymbol{\lambda}_i}(\boldsymbol{x})\Delta\lambda_{ji}^k$. The second line Taylor expands $\exp(-\Gamma_{ji}\frac{1}{2}\beta\Delta E_{ji}t)$ about $t = 0$.

We now simplify (G.29) using linear-response theory, effectively assuming that the most recent chemical transition is the dominant contribution to the present mechanical distribution (Ch. 7). The true 'initial' distribution over mechanical degrees of freedom following the previous chemical transition $\boldsymbol{\lambda}_\ell \to \boldsymbol{\lambda}_i$ is the (nonequilibrium) switching-state distribution $p_{i\ell}^{\text{sw}}(\boldsymbol{x})$. We approximate the difference between the mean conjugate force during relaxation from the previous nonequilibrium switching-state distribution $p_{i\ell}^{\text{sw}}(\boldsymbol{x})$ and at the nonequilibrium switching-state distribution $p_{ji}^{\text{sw}}(\boldsymbol{x},t)$ (that enters into the exact nonequilibrium excess work $\langle \beta W_{\text{ex}}^{\text{neq}} \rangle_{ji}$), by the difference between the mean conjugate force during relaxation from the previous equilibrium distribution $\pi_\ell(\boldsymbol{x})$ and at the current equilibrium distribution $\pi_i(\boldsymbol{x})$. Due to the symmetries of the model, these two pairs of distributions should differ by similar amounts. This substitution of one mean excess conjugate force for another is accurate when the conjugate forces $f_{\boldsymbol{x}|\boldsymbol{\lambda}_i}$ are approximately linear in $\boldsymbol{x}$ for all $\boldsymbol{x}$ with significant probability in $p_{\text{neq}}(\boldsymbol{x},t)$, $p_{ji}^{\text{sw}}(\boldsymbol{x},t)$, and the intervening relaxation. This approximation is trivially satisfied for a harmonic confining potential (such as in Sec. 10.4.1) where conjugate forces are always linear, and approximately satisfied for more general cases in the small-$\Delta\boldsymbol{\lambda}$ limit.

First-order Taylor expanding the mechanical potential $E(\boldsymbol{x}|\boldsymbol{\lambda}_\ell)$ around $\boldsymbol{\lambda}_\ell = \boldsymbol{\lambda}_i$ (similar to (G.27)) gives an approximate form for the previous equilibrium distribution

$$\pi_\ell(\boldsymbol{x}) = \frac{1}{\mathcal{Z}_\ell}\pi_i(\boldsymbol{x})e^{-\beta\Delta E_{i\ell}} \ . \tag{G.30}$$

Thus, using $\pi_\ell(\boldsymbol{x})$ as the initial condition for $p_{\text{neq}}(\boldsymbol{x},t)$ at $t=0$, we use linear-response theory to approximate the spatial integral in the nth term of (G.29b) in terms of correlation functions [110]:

$$\begin{aligned} I_k^{(n)}(t) &\equiv \int g_{\boldsymbol{\lambda}_i}^{(n)}(\boldsymbol{x})p_{\text{neq}}(\boldsymbol{x},t)e^{-\frac{1}{2}\beta\Delta E_{ji}}\,\mathrm{d}\boldsymbol{x} && \text{(G.31a)}\\ &= \frac{1}{\mathcal{Z}_\ell}\int g_{\boldsymbol{\lambda}_i}^{(n)}(\boldsymbol{x})\pi_i\, e^{-\beta\Delta E_{i\ell}}e^{\frac{1}{2}\beta\Delta E_{ji}}\,\mathrm{d}\boldsymbol{x} && \text{(G.31b)}\\ &\approx \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{\boldsymbol{\lambda}_i} - \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}(t)\beta\Delta E_{i\ell}\right\rangle_{\boldsymbol{\lambda}_i} + \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{\boldsymbol{\lambda}_i}\langle\beta\Delta E_{i\ell}\rangle_{\boldsymbol{\lambda}_i} + \frac{1}{2}\left\langle g_{\boldsymbol{\lambda}_i}^{(n)}(t)\beta\Delta E_{ji}\right\rangle_{\boldsymbol{\lambda}_i} . && \text{(G.31c)} \end{aligned}$$

Here the angle brackets $\langle\cdots\rangle_{\boldsymbol{\lambda}_i}$ indicate an average over the equilibrium distribution $\pi_i(\boldsymbol{x})$, and $g_{\boldsymbol{\lambda}_i}^{(n)}(\boldsymbol{x}) \equiv f_k|_{\boldsymbol{\lambda}_i}(\boldsymbol{x})\,(\beta\Delta E_{ji})^n$.

To calculate the nonequilibrium excess work beyond the TSS excess work outlined in Sec. 10.3.1 we subtract the linear-response approximation to the TSS excess work from (G.31c). The approximate TSS excess work, within the linear-response approximation of $g_{\boldsymbol{\lambda}}^{(n)}(\boldsymbol{x})$ averaged over the equilibrium switching-state distribution $p_{ji}^{\text{sw}}(\boldsymbol{x})$, is

$$\begin{aligned} \langle g_{\boldsymbol{\lambda}_i}^{(n)}\rangle_{ji}^{\text{sw}} &= \frac{1}{\mathcal{Z}_{ji}^{\text{sw}}}\int g_{\boldsymbol{\lambda}_i}^{(n)}(\boldsymbol{x})p_{ji}^{\text{sw}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} && \text{(G.32a)}\\ &\approx \frac{1}{\mathcal{Z}_{ji}^{\text{sw}}}\int g_{\boldsymbol{\lambda}_i}^{(n)}(\boldsymbol{x})\pi_i\left(1+\tfrac{1}{2}\beta f_k\Delta\lambda_{ji}^k\right)\mathrm{d}\boldsymbol{x} && \text{(G.32b)}\\ &= \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{\boldsymbol{\lambda}_i} + \frac{1}{2}\left\langle g_{\boldsymbol{\lambda}_i}^{(n)}(t)\beta\Delta E_{ji}\right\rangle_{\boldsymbol{\lambda}_i} , && \text{(G.32c)} \end{aligned}$$

which simplifies $I_k^{(n)}(t)$ to

$$I_k^{(n)}(t) = \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{ji}^{\text{sw}} - \beta\left[\left\langle g_{\boldsymbol{\lambda}_i}^{(n)}(t)\Delta E_{i\ell}\right\rangle_{\boldsymbol{\lambda}_i} + \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{\boldsymbol{\lambda}_i}\langle\Delta E_{i\ell}\rangle_{\boldsymbol{\lambda}_i}\right] . \tag{G.33}$$

The first RHS term approximates the TSS work, so subtracting $\langle g_{\boldsymbol{\lambda}_i}^{(n)}\rangle_{ji}^{\text{sw}}$ from both sides gives

$$\begin{aligned} \delta I_k^{(n)}(t) &\equiv I_k^{(n)}(t) - \left\langle g_{\boldsymbol{\lambda}_i}^{(n)}\right\rangle_{ji}^{\text{sw}} && \text{(G.34a)}\\ &= -\frac{\beta^{n+1}}{2^n}\Delta\lambda_{i\ell}^{k'}\left(\Delta\lambda_{ji}^{k'}\right)^n\langle\delta f_k^{n+1}(t)\delta f_{k'}(0)\rangle_{\boldsymbol{\lambda}_i} . && \text{(G.34b)} \end{aligned}$$

Beyond the leading-order ($n=0$) contribution, every term is $\mathcal{O}(\Delta\boldsymbol{\lambda}^2)$ or higher, hence beyond the approximation made in (G.33).

In (G.29), when the relaxation rates of the force autocovariances $\langle \delta f_k(0) \delta f_{k'}(t) \rangle_{\boldsymbol{\lambda}_i}$ are significantly larger than Γ_{ji}, the Taylor series is well-approximated by its leading term:

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (\Gamma_{ji} t)^n \int g^{(n)}_{\boldsymbol{\lambda}_i}(t) p_{\mathrm{neq}}(\boldsymbol{x}, t) e^{-\frac{1}{2}\beta \Delta E_{ji}} \, \mathrm{d}\boldsymbol{x} \approx -\beta \Delta \lambda^k_{i\ell} \langle \delta f_k(t) \delta f_{k'}(0) \rangle_{\boldsymbol{\lambda}_i} + \left\langle g^{(0)}_{\boldsymbol{\lambda}_i} \right\rangle^{\mathrm{sw}}_{ji} \ . \tag{G.35}$$

This approximation amounts to a statement of timescale separation between the chemical and mechanical dynamics, when the mechanical degrees of freedom relax significantly faster than the chemical-state dynamics. For instance, if the force autocovariance decays exponentially, $\langle \delta f_k(t) \delta f_{k'}(0) \rangle_{\boldsymbol{\lambda}_i} \propto \exp(-k_{\mathrm{relax}} t)$, then this approximation (G.34b) holds when $k_{\mathrm{relax}} \gg \Gamma_{ji}$.

Substituting (G.35) in the work (G.29) and subtracting the linear-response approximation to the TSS excess work $\langle g^{(0)}_{\boldsymbol{\lambda}_i} \rangle_{\boldsymbol{\lambda}_i}$ from both sides gives the nonequilibrium excess work

$$\langle W^{\mathrm{neq}}_{\mathrm{ex}} \rangle_{ji|\ell} \approx \beta \Gamma_{ji} \Delta \lambda^k_{ji} \Delta \lambda^{k'}_{i\ell} \int_0^\infty \langle \delta f_k(t) \delta f_{k'}(0) \rangle_{\boldsymbol{\lambda}_i} e^{-\Gamma_{ji} t} \, \mathrm{d}t \ , \tag{G.36}$$

which is the excess work required for the $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$ transition (given the previous state was $\boldsymbol{\lambda}_\ell$ prior to $\boldsymbol{\lambda}_i$), beyond the work required in the TSS limit. The average $\langle \cdots \rangle_{ji|\ell}$ is conditioned on the previous chemical state $\boldsymbol{\lambda}_\ell$. Once again, we expand the exponential as $\exp(-\Gamma_{ji} t) \approx 1$ (based on the same approximation simplifying the Taylor series in (G.35)), simplifying the nonequilibrium excess work (G.36) to

$$\langle W^{\mathrm{neq}}_{\mathrm{ex}} \rangle_{ji|\ell} \approx \Gamma_{ji} \Delta \lambda^k_{ji} \Delta \lambda^{k'}_{i\ell} \beta \int_0^\infty \langle \delta f_k(t) \delta f_{k'}(0) \rangle_{\boldsymbol{\lambda}_i} \mathrm{d}t \tag{G.37a}$$

$$= \Gamma_{ji} \Delta \lambda^k_{ji} \Delta \lambda^{k'}_{i\ell} \zeta_{kk'}(\boldsymbol{\lambda}_i) \ , \tag{G.37b}$$

where $\zeta_{kk'}(\boldsymbol{\lambda}) \equiv \beta \int_0^\infty \langle \delta f_k(0) \delta f_{k'}(t) \rangle_{\boldsymbol{\lambda}} \mathrm{d}t$ is the *generalized friction tensor* originally derived for continuous, deterministic control [109].

This approximation depends on three chemical states: $\boldsymbol{\lambda}_\ell$, $\boldsymbol{\lambda}_i$, and $\boldsymbol{\lambda}_j$. To simplify to be solely a function of the transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, we average over all previous chemical states $\boldsymbol{\lambda}_\ell$, with each term weighted by $P_\ell V_{i\ell} / V_{i*}$, the coarse-grained transition rate of $\boldsymbol{\lambda}_\ell \to \boldsymbol{\lambda}_i$ divided by the total entry rate into state $\boldsymbol{\lambda}_i$: $V_{i*} \equiv \sum_s P_s V_{is}$. Thus, the average excess work for transition $\boldsymbol{\lambda}_i \to \boldsymbol{\lambda}_j$, averaged over dwell-time fluctuations and previous states, within the linear-response regime is

$$\langle \beta W^{\mathrm{neq}}_{\mathrm{ex}} \rangle_{ji} \approx \beta \Gamma_{ji} \zeta_{kk'}(\boldsymbol{\lambda}_i) \Delta \lambda^k_{ji} \frac{\sum_s P_s V_{is} \Delta \lambda^{k'}_{is}}{V_{i*}} \ . \tag{G.38}$$

The corresponding approximate nonequilibrium excess power is

$$\langle \beta \mathcal{P}_{\rm ex}^{\rm neq} \rangle_{\Lambda \to X} \approx \sum_{ji} P_i \langle \beta \mathcal{P}_{\rm ex}^{\rm neq} \rangle_{ji} \tag{G.39a}$$

$$= \sum_{ji} P_i V_{ji} \langle \beta W_{\rm ex}^{\rm neq} \rangle_{ji} \tag{G.39b}$$

$$\approx \sum_{ji} P_i \beta V_{ji} \Gamma_{ji} \zeta_{kk'}(\boldsymbol{\lambda}_i) \Delta\lambda_{ji}^{k} \frac{\sum_s P_s V_{is} \Delta\lambda_{is}^{k'}}{V_{i*}} \ . \tag{G.39c}$$

All terms in (G.39) can be determined from coarse-grained observations, aside from the generalized friction tensor $\zeta_{kk'}(\boldsymbol{\lambda})$, a phenomenological quantity determined from conditional equilibrium measurements of the mechanical degrees of freedom.

G.3 Simulation details: linear-transport motor

In the model for a molecular transport motor (Sec. 10.4.1), simulations are implemented using the methods presented in App. A.3, and we measure both the excess work and heat dissipation. The nonequilibrium excess work (10.24) equals the sum of changes $\Delta E_{ji} = E(x_t|\lambda_j) - E(x_t|\lambda_i)$ in mechanical state energy during chemical transitions. Alternatively, the heat dissipated into the environment equals the sum of heats $Q_{\Delta t} \equiv E(x_{t+\Delta t}|\lambda_i) - E(x_t|\lambda_i)$ during mechanical state changes with the chemical state unchanging. In particular, for a trajectory with N_Λ chemical transitions $\lambda_0 \to \lambda_1 \to \cdots \to \lambda_{N_\Lambda}$, with each chemical state having a dwell time $\tau_0, \tau_1, \cdots, \tau_{N_\Lambda}$, and each transition occurring at times $t_{\lambda_{10}}, t_{\lambda_{21}}, \cdots, t_{\lambda_{N_\Lambda, N_{\Lambda-1}}}$, the excess work and heat per step are

$$\langle \beta W_{\rm ex}^{\rm neq} \rangle_{\Delta\lambda} = \frac{\beta W}{N_\Lambda} \tag{G.40a}$$

$$= \frac{\beta}{N_\Lambda} \sum_{i=0}^{N_\Lambda - 1} E(x_{t_{i+1,i}}|\lambda_{i+1}) - E(x_{t_{i+1,i}}|\lambda_i) \tag{G.40b}$$

$$\langle \beta Q \rangle = \frac{\beta Q}{N_\Lambda} \tag{G.40c}$$

$$= \frac{\beta}{N_\Lambda} \sum_{i=0}^{N_\Lambda - 1} \sum_{j=0}^{M_{i-1}} E(x_{t_{j+1}}|\lambda_i) - E(x_{t_j}|\lambda_i) \ , \tag{G.40d}$$

where $M_{i-1} = \tau_{i-1}/\Delta t$ is the number of time steps that the chemical state remained at λ_{i-1}. The work equals the nonequilibrium excess work in (G.40b) because, for this system the equilibrium free energy change is zero ($\Delta F_{ji} = 0$) and $\langle \beta W_{\rm ex}^{\rm TSS} \rangle_{ji} = 0$.

The heat flow into the reservoir equals the entropy production in the environment due to the mechanical system dynamics (F.2). Figure G.1 shows for the simulated system the excess work and negative heat per step, showing their equivalence. This implies that the excess work in this system is, indeed, equal to the entropy production in the environment due to the mechanical dynamics; however, as pointed out in Ch. 9 and Refs. [195, 199], for strongly coupled systems the entropy produced in the reservoir due to one subsystem is not constrained to be positive by the typical form of the second law.

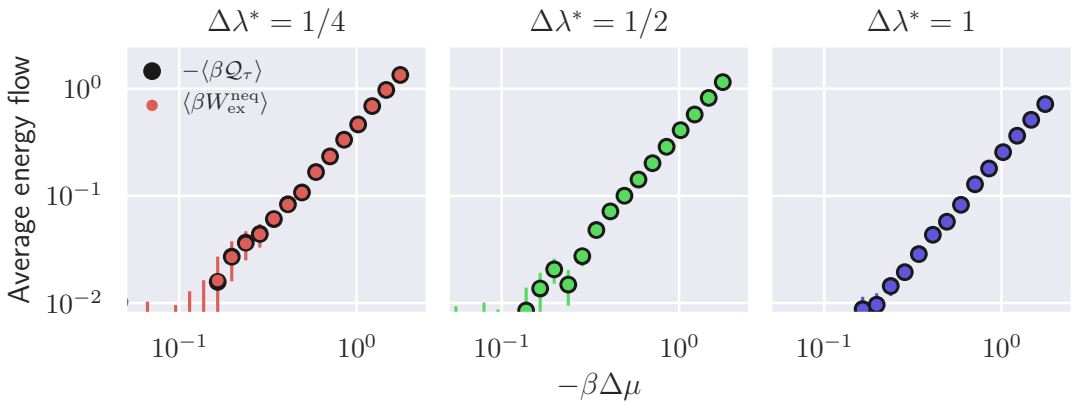

Figure G.1: **At steady state, excess work equals the negative heat.** Average excess work per step (coloured dots) $\langle\beta W_{\rm ex}\rangle_{\Delta\lambda}$ equals the heat flow (black dots) $-\langle\beta Q\rangle$ into the reservoir in between jumps, over the same range of chemical potential differences $-\beta\Delta\mu$ and control-parameter jump sizes $\Delta\lambda$ considered in Fig. 10.2.